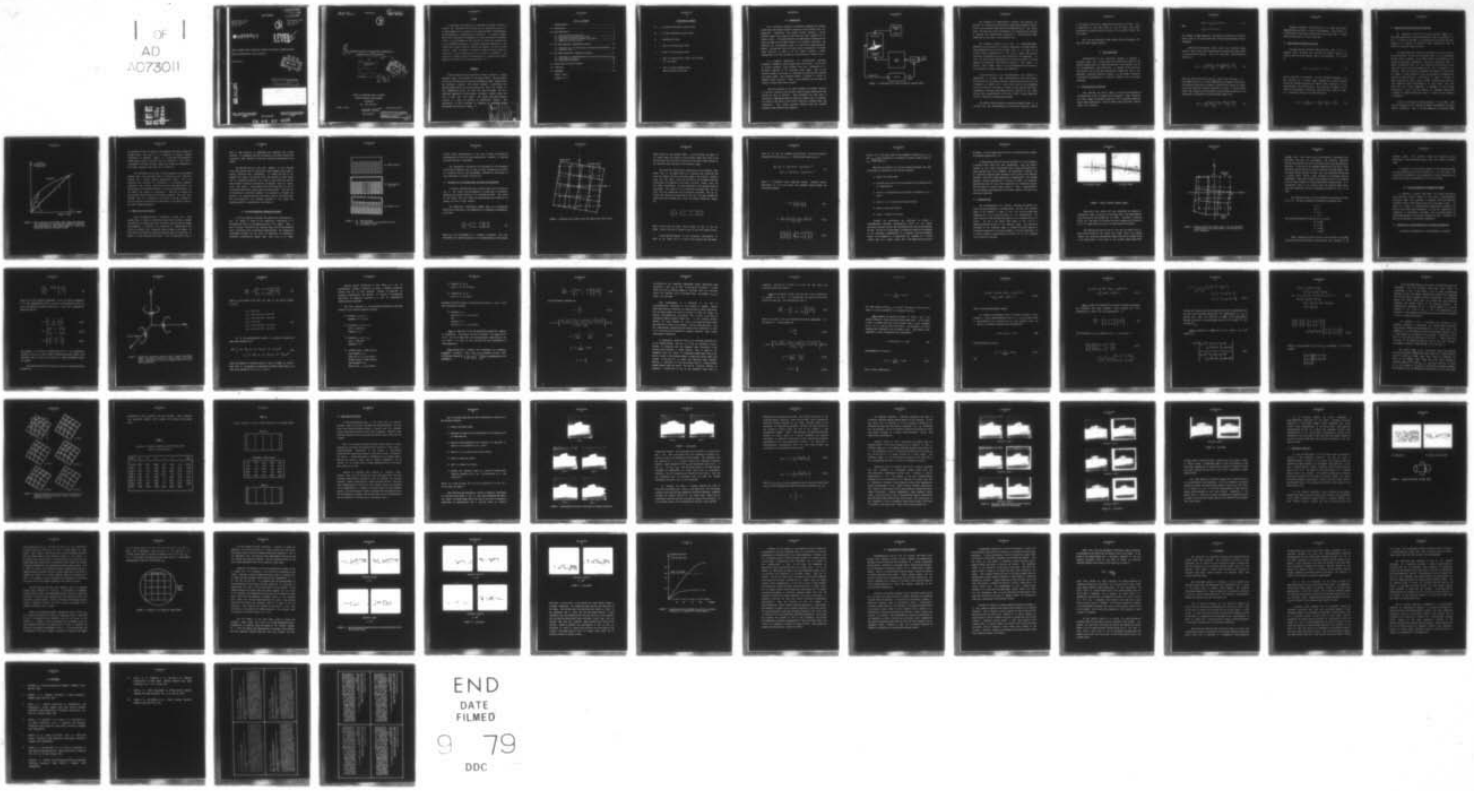
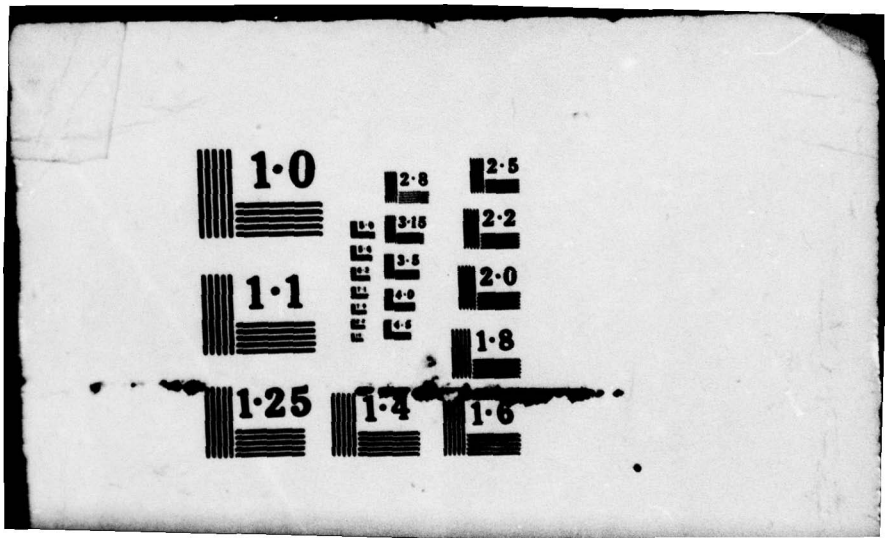
AD-A073 011 DEFENCE RESEARCH ESTABLISHMENT VALCARTIER (QUEBEC) F/G 17/8
AREA CORRELATION TRACKING WITH AUTOMATIC COMPENSATION FOR MAGNI--ETC(U)
JUL 79 C MUNTEANU

UNCLASSIFIED DREV-R-4140/79 NL

1 OF 1
AD
A073011



END
DATE
FILMED
9 79
DDC


1·0
1·1
1·25
2·8
3·15
3·5
1·4
2·5
2·2
2·0
1·8
1·6

UNCLASSIFIED

CRDV RAPPORT 4140/79
DOSSIER: 3621J-003
JUILLET 1979

DREV REPORT 4140/79
FILE: 3621J-003
JULY 1979

(3)

AD A073011

LEVEL II

AREA CORRELATION TRACKING WITH AUTOMATIC COMPENSATION FOR MAGNIFICATION AND ROTATION

C. Munteanu

Centre de Recherches pour la Défense
Defence Research Establishment
Valcartier, Québec

BUREAU - RECHERCHE ET DÉVELOPPEMENT
MINISTÈRE DE LA DÉFENSE NATIONALE
CANADA

RESEARCH AND DEVELOPMENT BRANCH
DEPARTMENT OF NATIONAL DEFENCE
CANADA

NON CLASSIFIÉ

CRDV R-4140/79
DOSSIER: 3621J-003

UNCLASSIFIED

DREV-R-4140/79
FILE: 3621J-003

(3)

AREA CORRELATION TRACKING WITH AUTOMATIC COMPENSATION FOR MAGNIFICATION AND ROTATION

by

C. Munteanu

(11) Jul 79

(12) 72p.

CENTRE DE RECHERCHES POUR LA DEFENSE
DEFENCE RESEARCH ESTABLISHMENT
VALCARTIER
Tel. (418) 844-4271

Quebec, Canada

July/juillet 1979

404945

NON CLASSIFIE

RESUME

La poursuite d'une cible par la corrélation d'images consiste à corréler une image étalon de la cible avec une scène prise sur le vif. Le grossissement et les rotations de la cible diminuent considérablement les performances de l'algorithme de corrélation. Le présent rapport décrit deux algorithmes qui permettent à la tête chercheuse de compenser l'effet du grossissement et des rotations à l'aide d'une lentille zoom et d'un système de balanciers. Un premier algorithme, applicable à des cibles fixes, évalue l'angle de rotation par rapport à la ligne de visée (roulis). Quant au second, il détermine l'angle de rotation selon les trois axes, ce qui est nécessaire pour une cible en mouvement. Une simulation réalisée à partir de photographies et d'images artificielles démontre l'efficacité de ces deux algorithmes. (NC)

ABSTRACT

Target tracking by area correlation involves matching a stored reference image of the target to live search scenery. The performance of the area correlator can be severely degraded by target magnification and three-axis rotation. This report presents 2 algorithms that estimate magnification and rotation angles, and allow the tracker to take compensating action via zoom lens and roll gimbal. The first algorithm, applicable to fixed targets, estimates rotation about the line of sight only (roll angle). The second algorithm estimates rotation in 3 axes, as required for maneuvering targets. The presentation of both algorithms is supported by simulation using photographs and artificial images. (U)

TABLE OF CONTENTS

RESUME/ABSTRACT. i

1.Ø INTRODUCTION 1

2.Ø AREA CORRELATION 4

2.1 Cross-Correlation Algorithm 4
2.2 Mean Absolute Difference Algorithm. 6
2.3 Sequential Similarity Detection Algorithm 7
2.4 Magnification and Rotation. 9

3.Ø THE TWO-DIMENSIONAL COMPENSATION PROBLEM 1Ø

3.1 Estimation of a Two-Dimensional Coordinate Transformation . 12
3.2 Simulation Test 17

4.Ø THE THREE-DIMENSIONAL COMPENSATION PROBLEM 21

4.1 Estimation of a Three-Dimensional Coordinate Transformation 21
4.2 Small-Angle Simulation. 39
4.3 Large-Angle Simulation. 48

5.Ø APPLICATION TO MISSILE GUIDANCE. 58

6.Ø CONCLUSION 61

7.Ø REFERENCES 64

TABLES I AND II

FIGURES 1 TO 14

DEFINITION OF SYMBOLS

Δx = horizontal translation (along x-axis)

Δy = vertical translation (along y-axis)

M = magnification ratio

α = angle of rotation about x-axis

β = angle of rotation about y-axis

γ = angle of rotation about z-axis (line of sight) i.e. roll angle

Γ = total roll angle between sensor axes and pitch and yaw axes

1.0 INTRODUCTION

Area correlation tracking is an effective technique for tracking targets that have neither prominent features nor high contrast with the background. Essentially, this method involves matching a stored reference image of the target to real-time scene data acquired by an imaging sensor. The image-matching is accomplished by superposing the reference image over the search scene at all possible positions, and computing the corresponding values of some kind of image similarity measure. The position within the search scene corresponding to maximum image similarity is taken to be the target's location. A basic configuration of an area correlation tracking system is shown in Fig. 1.

In a tracking application, the image-matching algorithm encounters problems of target magnification and rotation. In the field of missile guidance, these problems have been circumvented by a periodical updating of the reference target image. That is, new reference images of the target are repeatedly selected from the incoming live search scenes. This procedure, however, is subject to errors in computed target position. These errors accumulate and can cause the tracker to wander away from the target.

The work presented in this report addresses the problem directly by deriving algorithms that estimate relative target magnification and rotation as well as position. These estimates are then used in controlling compensating devices (zoom lens, gimbals, etc.) to maintain the target at the center of the field of view with a constant scale and orientation. This allows continued tracking using the original reference target image without updating.

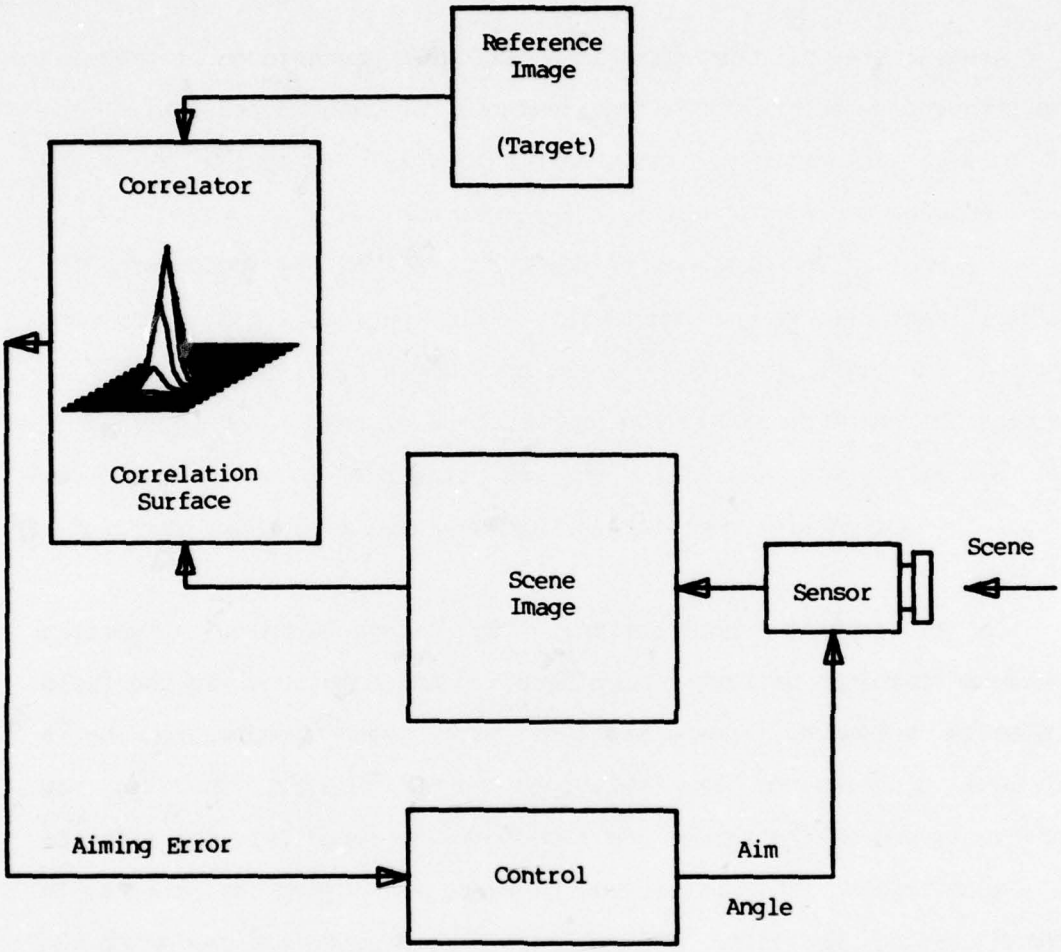


FIGURE 1 - Configuration for area correlation tracking system

The estimates of magnification, rotation, and position are obtained by estimating the coordinate transformation between a grid of points in the reference image and the corresponding grid in the search scene. The search grid that corresponds to the reference grid is formed by applying the image-matching algorithm to relatively small image subsections at each point of the reference grid.

The estimation problem can be two- or three-dimensional, depending on the exclusion or inclusion respectively of target rotation about perpendiculars to the line of sight. That is, if there is no rotation about perpendiculars to the line of sight, then the grid coordinate transformation can be fully described in two dimensions. If there is rotation about perpendiculars to the line of sight, then three dimensions are required. Both cases are treated in this report. In general, fixed targets require only a two-dimensional grid analysis, while maneuvering targets require three dimensions.

A major difficulty in the three-dimensional case consists in compensating for rotation about perpendiculars to the line of sight. While effects such as target translation, magnification and roll can be eliminated via gimbals and a zoom lens, there is no corresponding corrective action for the case of a target turning away from the sensor. Consequently, reference image updating will be required as the target rotates about perpendiculars to the line of sight. Target image scale and roll orientation are maintained constant.

The tracking system acquires an important capability when it is equipped with the three-dimensional grid analysis algorithm. Without

it, the system tracks only an image which can rotate out of view. With 3 dimensions, on the other hand, it will be shown that, within limitations, the system tracks the target as a true object rather than as an image.

This work was performed at DREV between March and September 1977 under PCN 21J03 "Imaging Seekers".

2.0 AREA CORRELATION

Image-matching by area correlation consists of scanning a reference image over a search image and recording the location of the maximum computed value of a similarity function. This function can be the two-dimensional cross-correlation, as implied by the term "area correlation", but this has been generalized to include the use of any meaningful similarity function or measure. For the reader's reference, 3 algorithms corresponding to 3 similarity measures will be briefly described.

2.1 Cross-Correlation Algorithm

The reference and search images are digitally represented as two-dimensional arrays of integers that correspond to image intensity. Let these images be denoted by matrices W and S (reference window and search scene respectively). Then the simple cross-correlation function (Refs. 1-3) is given by

$$R(u,v) = \sum_{ij}\sum S_{uv}(i,j)\, W(i,j) \qquad [1]$$

where

$$S_{uv}(i,j) \equiv S(u+i,v+j)$$

As a measure of image similarity, this function can easily be tricked in other areas of the search scene that correlate highly because of their brightness.

These false correlations, which result from nonuniform image contrast and dc level, can be eliminated by incorporating normalization in the correlation function. Hence the normalized cross-correlation function is

$$R(u,v) = \sum_{ij}\sum \left(\frac{S_{uv}(i,j) - \bar{S}_{uv}}{\sigma_{S_{uv}}} \right)\left(\frac{W(i,j) - \bar{W}}{\sigma_W} \right) \qquad [2]$$

where $\bar{S}_{uv}$, $\bar{W}$ are local means, and $\sigma^2_{S_{uv}}$, σ^2_W are local variances. At every spatial shift (u,v), before multiplying and summing (correlating), this function normalizes the images to identical mean values (dc levels) of zero and identical variances (contrasts) of unity. It can be readily shown that maximizing the normalized cross-correlation function [2] is equivalent to minimizing the normalized mean square error

$$R(u,v) = \sum_{ij}\sum \left(\frac{S_{uv}(i,j) - \bar{S}_{uv}}{\sigma_{S_{uv}}} - \frac{W(i,j) - \bar{W}}{\sigma_W} \right)^2 \qquad [3]$$

Although normalized cross-correlation is a robust and reliable image-matching algorithm, it is very time-consuming. In this report, it is used primarily with small reference and search areas after a coarse search has been done by a faster algorithm.

2.2 Mean Absolute Difference Algorithm

The mean absolute difference (MAD) algorithm (Refs. 4-5) is a somewhat faster algorithm that replaces the product operation of cross-correlation by the absolute value of the difference. It is given by

$$R(u,v) = \sum_{ij}\sum \left| S_{uv}(i,j) - W(i,j) \right| \qquad [4]$$

and is a function to be minimized. For most arithmetic processors, the MAD algorithm has a notable speed advantage over cross-correlation because the combined subtraction and absolute-value operation is less time-consuming than multiplication. This algorithm is, however, not as robust in the presence of significant image noise and distortion (Ref. 4). A normalized version of the MAD algorithm is

$$R(u,v) = \sum_{ij}\sum \left| \left(S_{uv}(i,j) - \overline{S_{uv}} \right) - \left(W(i,j) - \overline{W} \right) \right| \qquad [5]$$

2.3 Sequential Similarity Detection Algorithm

The sequential similarity detection algorithm (SSDA) is a high-speed image-matching algorithm presented in Ref. 6 and analyzed in Ref. 7. The basis of the SSDA is the observation that a great deal less computation is required to determine image dissimilarity than is required to determine image similarity.

The image-comparison criterion of the SSDA is the same as that of the MAD algorithm: the sum of the absolute values of the differences between picture elements. In the SSDA, however, this summation is not necessarily performed to completion. The picture elements are taken in random nonrepeating order (so as to maximize new information at each step) and at every step of the summation, the sum is subjected to a threshold. When the threshold is exceeded, the comparison is halted and the number of steps required to cross the threshold is recorded as the function of image similarity. This function is small in regions of image dissimilarity where picture element differences are large and their summation crosses the threshold at an early stage. This also means that very little computation is expended in areas of image dissimilarity. Since the vast majority of candidate match-points correspond to image dissimilarity, there is a drastic overall reduction in computation.

Figure 2 illustrates the typical operation of the SSDA. Note that the threshold is not a constant throughout the summation of picture element differences. This is so that comparison of dissimilar images

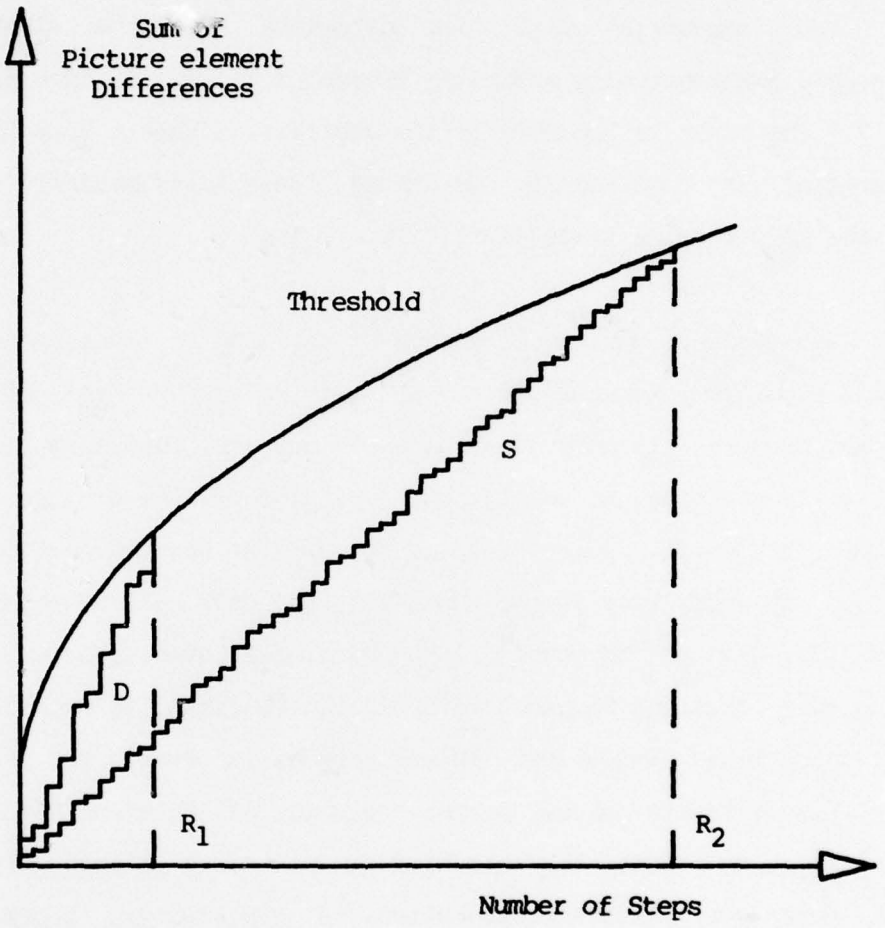


FIGURE 2 - Typical operation of the SSDA. When images are dissimilar (D), cumulative error grows quickly, giving low value of similarity function. With similar images (S), error grows slowly and similarity function is high.

can terminate as early as possible, and comparison of similar images can proceed fully. The curve D is a summation of picture element differences of dissimilar images. It is seen that the threshold is crossed early, giving a low value (R_1) of similarity. The curve S corresponds to the case of similar images and crosses the threshold at a late stage, yielding a high value (R_2) of similarity.

The performance of the SSDA is highly dependent on the threshold sequence used. A low threshold leads to a sharp peak in the similarity function and a greatly reduced computational effort. But a low threshold also increases the probability of terminating summation prematurely and, thereby, missing the correct image match-point. The determination of an optimum threshold is a difficult problem because of the SSDA's high degree of nonlinearity. Reference 6 presents two heuristically derived families of threshold sequences that work well although they are not designed to be optimal in any way. Reference 7 gives an analysis that predicts the performance of the SSDA for these or any other threshold sequences.

2.4 Magnification and Rotation

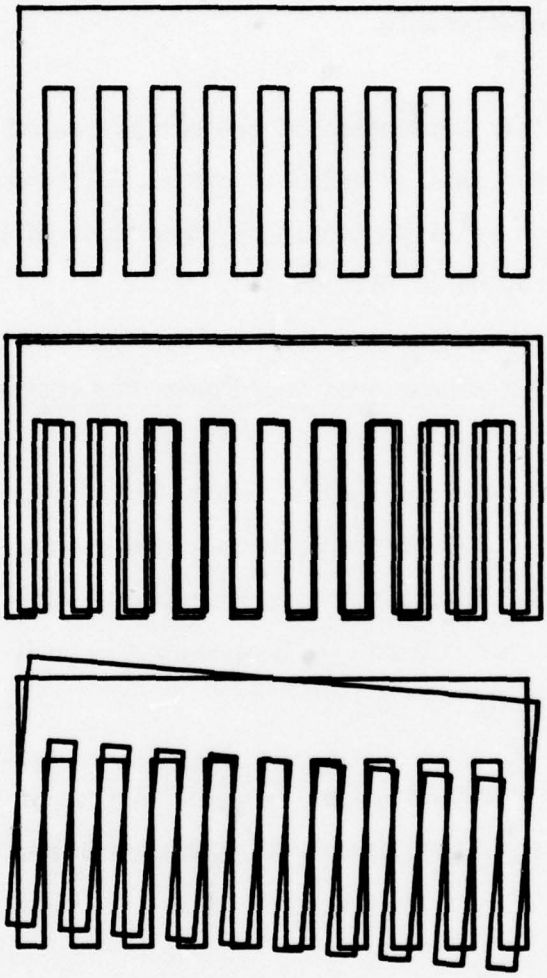
Since the image-matching is essentially a search for a best superposition of images, it should be evident that this task is impeded by differences in magnification and orientation angle between the images being compared. In addition, the sensitivity to magnification and rotation increases with increasing detail present in the images. In other words, the overlap of images of different scale or orientation is weakest in areas containing fine detail. This is illustrated in Fig. 3

where a comb structure is superposed with magnified and rotated versions. It is apparent that the fine detail of the teeth retains less correlation under geometric distortion than the broad feature of the handle.

The important point to note here, however, is that although entire comb pictures may not overlap well, small segments can be made to overlap relatively well individually. In Figures 3(b) and (c), for example, the 4 teeth in the center of the comb overlap to a good extent with no translation required. The 3 teeth at each end of the comb can also be made to overlap if translations are applied. Thus, the result, rather than being a single position for the second comb, is a set of translations describing a rubber-sheet stretching effect. From this set of translations of individual segments, it is possible to estimate the position, the magnification and the rotation angle of the second comb. This is the essence of the technique presented in this report for estimating position, magnification, and rotation.

3.0 THE TWO-DIMENSIONAL COMPENSATION PROBLEM

In the grid analysis technique, the geometrical relationship of a pair of images is given by the geometrical relationship of a representative subset of corresponding points in the images. Thus, a grid of points is defined in the reference image, and the corresponding grid in the search scene is located by image-matching. The geometrical relationship between the two grids is determined by estimating a coordinate transformation between them. When there is no target



(a) Comb structure

(b) Magnification of 1.05

(c) Rotation of 5°

FIGURE 3 - (a) Comb structure
(b) in superposition with magnified version
(c) and rotated version.

rotation about perpendiculars to the line of sight, the problem is two-dimensional in that only two coordinates are required to describe the grids and their relationship.

The mathematical derivation of the estimation of the coordinate transformation between a pair of grids is given, followed by the results of a simulation test on still photographs. Extension of the problem to three dimensions is treated in Section 4.0.

3.1 Estimation of a Two-Dimensional Coordinate Transformation

A 25-point reference grid and a 25-point search grid are shown in Fig. 4. The search grid corresponds to a target that is translated, magnified, and rotated relative to its reference attitude. The points of the search grid are found by image-matching and, therefore, are subject to errors in their location.

The geometrical relationship between pairs of corresponding points can be described in two dimensions by a coordinate transformation of the form

$$\begin{bmatrix} u \\ v \end{bmatrix} = M \begin{bmatrix} \cos\gamma & \sin\gamma \\ -\sin\gamma & \cos\gamma \end{bmatrix} \begin{bmatrix} x-\Delta x \\ y-\Delta y \end{bmatrix} \qquad [6]$$

where (x,y) are coordinates of a reference grid-point, (u,v) are coordinates of a search grid-point, M is the magnification of the search

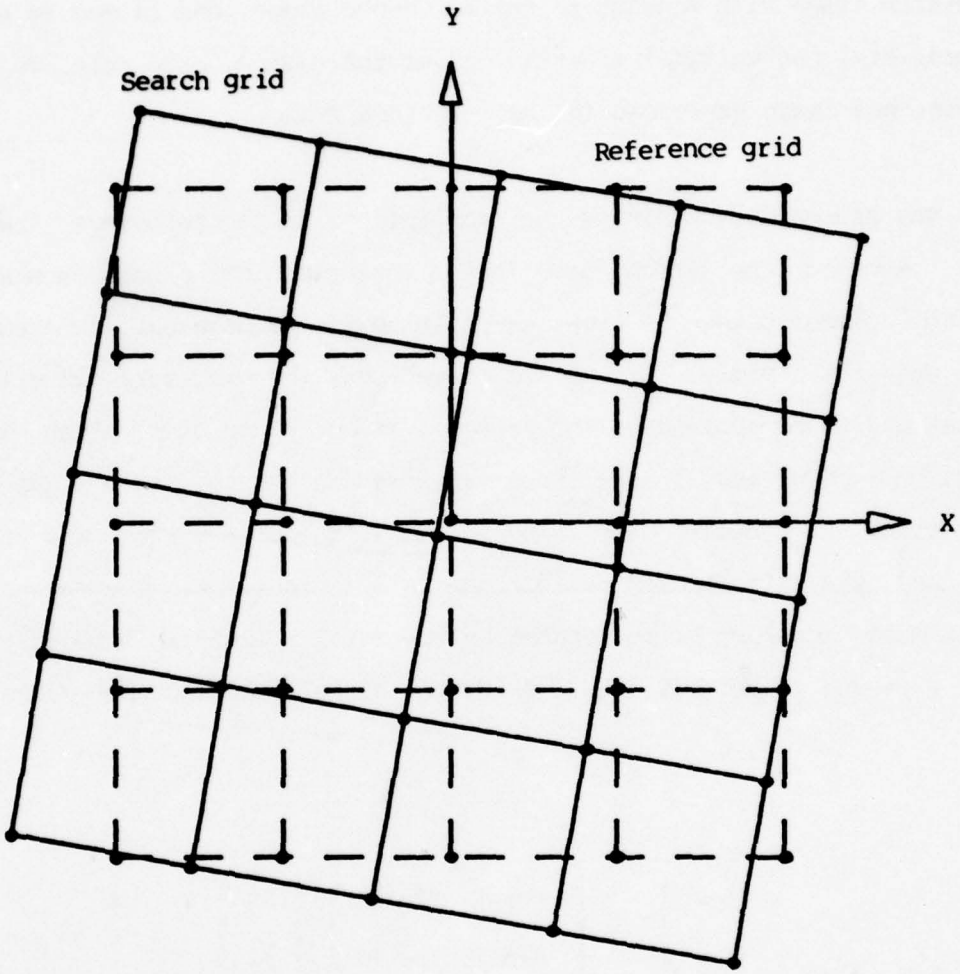


FIGURE 4 - Reference grid (broken lines) and search grid (solid lines)

image relative to the reference image, γ is the clockwise roll angle of the search image with respect to the reference image, and Δx and Δy are the horizontal and vertical translations of the search image relative to the reference image expressed in the reference frame.

The reason for obtaining the translations in the reference frame rather than in the search frame is the intention that a zoom lens and inner roll gimbal bring the live search image to the original reference scale and roll angle. Given this correction, the necessary centering translations to be applied to the sensor (via outer tracker pitch and yaw gimbal adjustments) are those expressed in the reference frame of coordinates. As a result of roll corrections, the sensor axes and the pitch and yaw gimbal axes are not generally coincident. The (Δx, Δy) signals must therefore be prerotated by the total roll angle between the sensor axes and the gimbal axes. Thus, the signals feeding the gimbals are

$$\begin{bmatrix} \Delta x' \\ \Delta y' \end{bmatrix} = \begin{bmatrix} \cos \Gamma & \sin \Gamma \\ -\sin \Gamma & \cos \Gamma \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$$

where Γ is the total roll angle. The $\Delta x'$ signal is sent to the yaw gimbal control, and the $\Delta y'$ signal is sent to the pitch gimbal control.

A least squares estimate of the coordinate transformation is given by the values of M, γ, Δx and Δy that minimize the mean square

error of [6] for the ensemble of grid points. Let the grid data be arranged as N-vectors $\underline{x}$, $\underline{y}$, $\underline{u}$, $\underline{v}$. Then the mean square error is

$$\begin{aligned} MSE = \frac{1}{N} \| \underline{u} - M(\underline{x}-\Delta x)\cos\gamma - M(\underline{y}-\Delta y)\sin\gamma \|^2 \\ + \frac{1}{N} \| \underline{v} + M(\underline{x}-\Delta x)\sin\gamma - M(\underline{y}-\Delta y)\cos\gamma \|^2 \end{aligned} \qquad [7]$$

where $\| \ \|^2$ indicates vector magnitude squared. Equating partial derivatives of [7] to zero yields, after somewhat lengthy algebra, the following estimates:

$$\gamma = \tan^{-1}\left(\frac{C_{uy} - C_{vx}}{C_{ux} + C_{vy}}\right) \qquad [8.1]$$

$$M = \frac{(C_{ux} + C_{vy})\cos\gamma + (C_{uy} - C_{vx})\sin\gamma}{C_{xx} + C_{yy}} \qquad [8.2]$$

$$\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} - \frac{1}{M}\begin{bmatrix} \cos\gamma & \sin\gamma \\ -\sin\gamma & \cos\gamma \end{bmatrix}^{-1}\begin{bmatrix} \bar{u} \\ \bar{v} \end{bmatrix} \qquad [8.3]$$

where $\bar{x}$, $\bar{y}$, $\bar{u}$, $\bar{v}$ are mean values of the elements of vectors $\underline{x}$, $\underline{y}$, $\underline{u}$, $\underline{v}$, and C_{ux} is the covariance of the elements of vectors $\underline{u}$ and $\underline{x}$; that is, $C_{ux} = \overline{(u-\bar{u})\ (x-\bar{x})}$ etc.

What will be referred to as the 2-D tracking algorithm can then be described as repetition of the following sequence:

1) Acquire the search scene

2) Determine the search grid corresponding to the reference grid by image-matching.

3) Perform a two-dimensional grid analysis to estimate Δx, Δy, γ and M.

4) Send $\Delta x'$, $\Delta y'$ to sensor pitch and yaw control.

5) Send M to zoom lens control.

6) Send γ to sensor roll control.

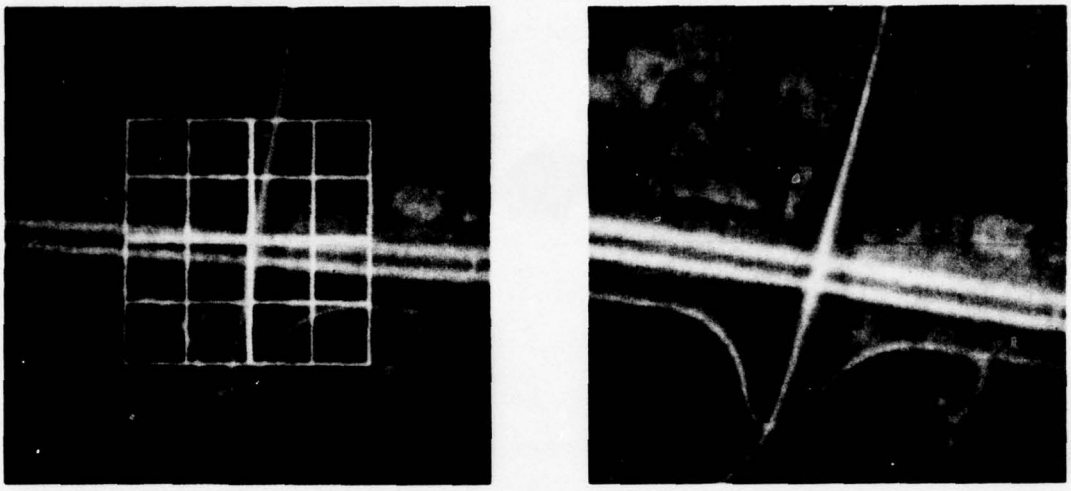
Although two point-pairs are sufficient to define a two-dimensional coordinate transformation, errors in the search grid-point locations dictate that considerably more than two grid-points be used. The use of a large number of grid-points improves the estimate of the coordinate transformation by increasing the error-averaging effect. As an example, if N grid-points are arranged on a circle (rather than on a square array) and if the search grid errors are

zero-mean, it can be shown that the error in the magnification estimate is inversely proportional to $\sqrt{N}$.

In obtaining the search grid point locations, it is not necessary to perform a wide search for each subreference. Once the general location of the target is determined from a single wide search with a large reference (32 x 32, for example), the subreference locations may be found with relatively small searches in the vicinity of the expected locations. The size of these subsearches is determined by the maximum expected relative magnification ratio and roll angle. For reasonable magnification and roll rates, these subsearches are considerably smaller than the search for overall target position. Hence, computationally, the grid analysis algorithm does not add great expense to the basic area correlation algorithm.

3.2 Simulation Test

The two-dimensional grid analysis technique was applied to an aerial infrared photograph of a highway scene. The 128 x 128 reference image of the highway is shown in Fig. 5(a) with the target area segmented into a 4 x 4 array of sixteen 16 x 16 subreferences. The center points of these subreferences define a 4 x 4 reference grid of 16 points. The 128 x 128 search scene shown in Fig. 5(b) is a computer-processed version of the reference image. The operations performed on the reference image to produce the search image are a horizontal translation of 2 picture elements, a vertical translation of 5 picture elements, a magnification by a factor of 1.1, and a clockwise roll rotation of 5 degrees.



(a) Reference image (b) Search image

FIGURE 5 - Aerial infrared highway imagery

Each point of the search grid was determined by matching a subreference image to an area of the search scene. The image-matching algorithm used for this purpose was the SSDA. Fractional coordinate values for each point of the search grid were computed as the centroid of a 5 x 5 region about the peak of the SSDA similarity function.

The reference and search grids for this test are shown in Fig. 6. Note that two points of the search grid (upper right) are missing because the associated subreference images could not be uniquely found in the search scene. In the case of the extreme upper-right-corner

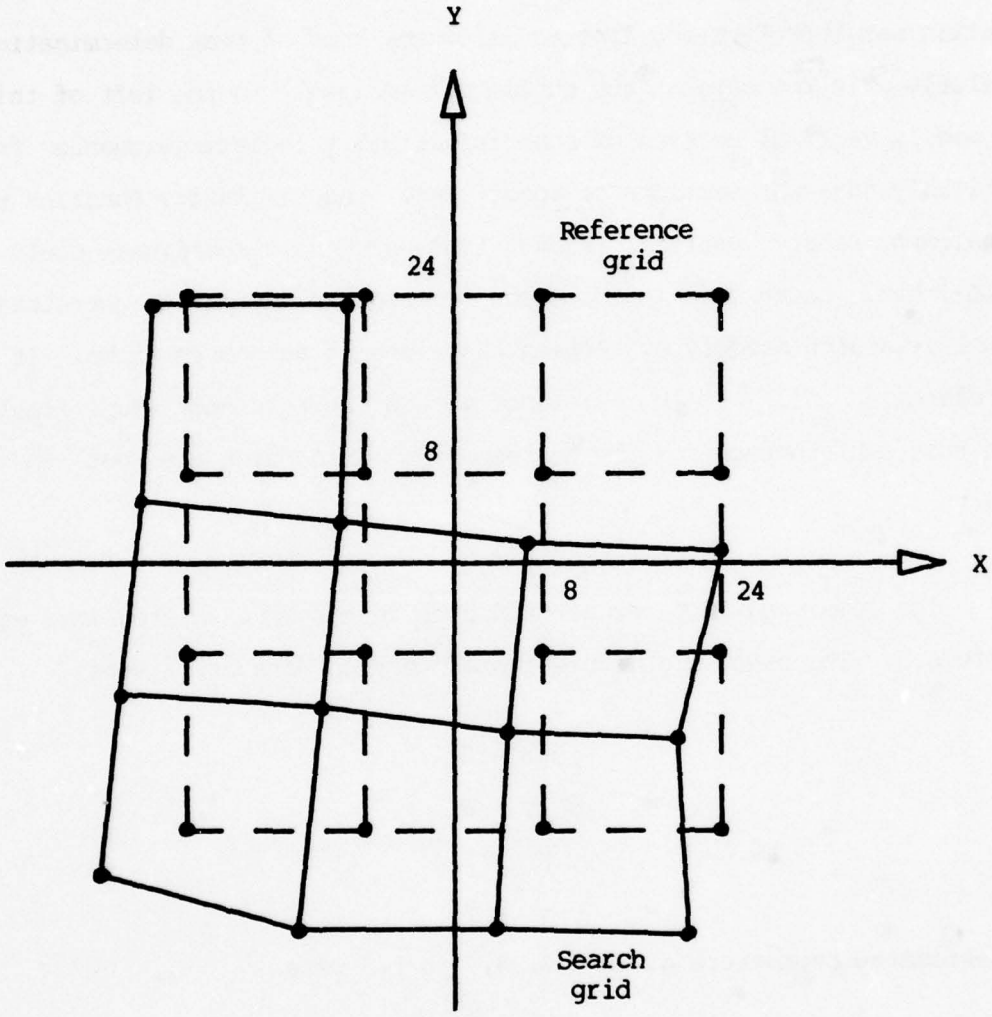


FIGURE 6 - Highway reference and search grids. The dots represent the locations of the center-points of the reference and search subimages.

subimage (Fig. 5(a)), there is very little content to distinguish that particular area from its surroundings. The resulting similarity function was therefore too flat to allow any kind of peak determination. Similarly, in the case of the subimage immediately to the left of this, the nearly vertical section of road is virtually indistinguishable from vertically adjacent sections of road. Here, the similarity function was a uniform nearly vertical ridge from which no y-coordinate could be established. Cases such as these can be weeded out prior to correlation with live search imagery by testing subreference autocorrelation. If a subreference fails to be distinct within the reference image itself, then that subreference and the corresponding grid-point are not to be used.

The remaining 14 pairs of grid coordinates were entered into eqs. 8.1 to 8.3. The exact coordinate transformation parameters were

$$\Delta x = 2$$
$$\Delta y = 5$$
$$M = 1.1$$
$$\gamma = 5°$$

The estimated parameters given eqs. 8.1 to 8.3 were

$$\Delta x = 2.320$$
$$\Delta y = 5.128$$
$$M = 1.093$$
$$\gamma = 5.276°$$

These estimates are used to direct a zoom lens and a roll gimbal to bring the search scene back to the scale and roll attitude of the

reference image. The estimates reduce the orientation errors sufficiently to allow continued tracking with the original reference image.

A discussion of the practical aspects of this technique applied to missile guidance will be left to Section 5.0. At this point, the problem will be extended to three dimensions to include consideration of target rotation about perpendiculars to the line of sight.

4.0 THE THREE-DIMENSIONAL COMPENSATION PROBLEM

To determine a general description of a target's attitude in space (including rotation angles in three axes), three-dimensional information about the grid-points is required. The third coordinate represents relative range (depth) along the line-of-sight (z-axis). An image of a target is a two-dimensional representation, however, that does not give direct range information. To avoid the complication of integrating a range-finding instrument into the tracking system, the algorithm presented in this section estimates the range information as well as the coordinate transformation parameters.

4.1 Estimation of a Three-Dimensional Coordinate Transformation

A coordinate transformation in three dimensions is given by

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = MR_\gamma R_\beta R_\alpha \begin{bmatrix} x-\Delta x \\ y-\Delta y \\ z \end{bmatrix} \qquad [9]$$

where (x,y,z) are reference coordinates, (u,v,w) are search coordinates, M is the magnification factor, Δx and Δy are translations, and R_α, R_β and R_γ are matrices of rotation about the x, y and z axes respectively. These matrices are

$$R_\alpha = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{bmatrix} \qquad [10.1]$$

$$R_\beta = \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix} \qquad [10.2]$$

$$R_\gamma = \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad [10.3]$$

The geometry of rotation is illustrated in Fig. 7. It is required to estimate Δx, Δy, M, α, β and γ given a reference grid and a search grid described only by sets of (x,y) and (u,v). Range coordinates z and w are unknown.

The equations relating the given (x,y) and (u,v) grid data can be rewritten as

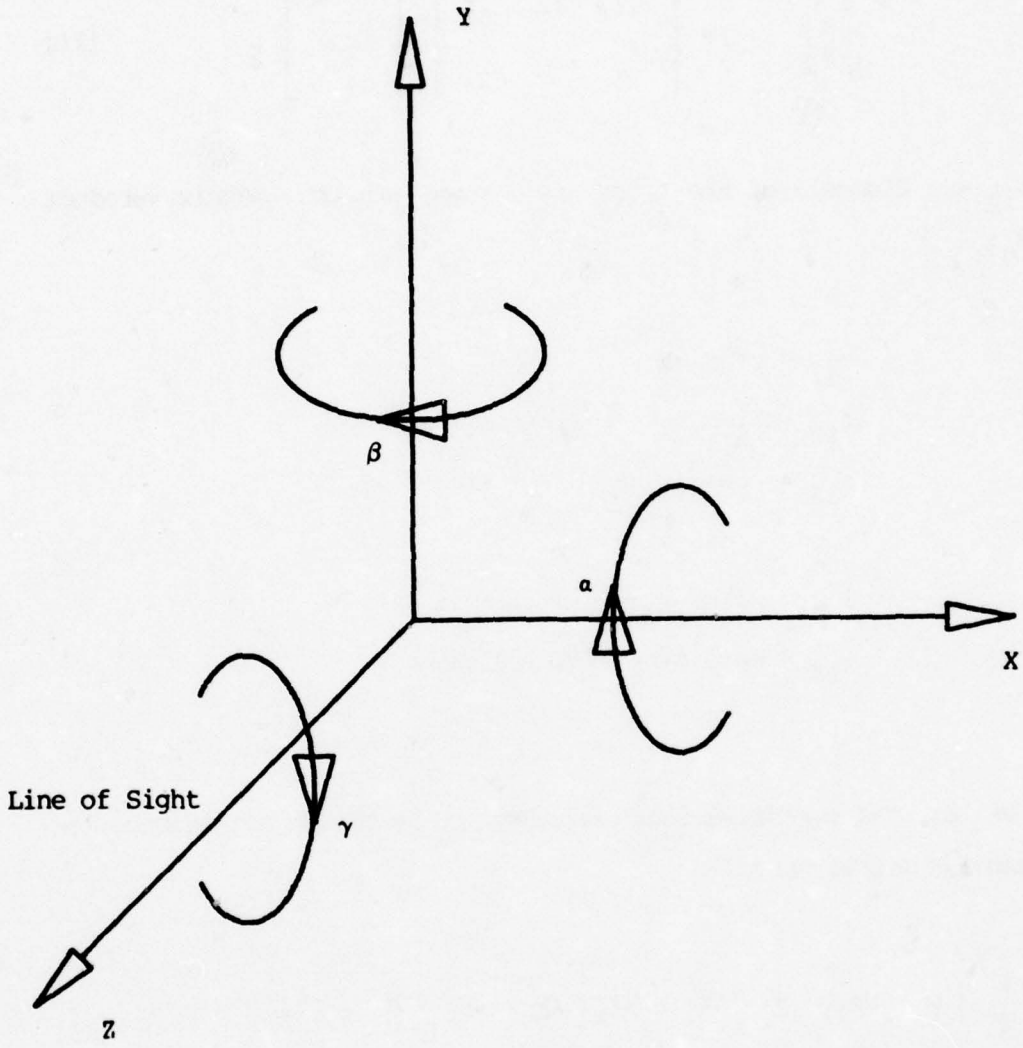


FIGURE 7 - Geometry of three-axis rotation. The x-y plane is the image plane. The z-axis is the line of sight. Angle γ is the roll angle. Translations Δx and Δy are measured along the x and y axes respectively.

$$\begin{bmatrix} u \\ v \end{bmatrix} = M \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \end{bmatrix} \begin{bmatrix} x-\Delta x \\ y-\Delta y \\ z \end{bmatrix} \qquad [11]$$

where r_{ij} are elements of the first two rows of the matrix product $R_\gamma R_\beta R_\alpha$:

$$\begin{aligned} r_{11} &= \cos\beta \cos\gamma \\ r_{12} &= \sin\alpha \sin\beta \cos\gamma + \cos\alpha \sin\gamma \\ r_{13} &= -\cos\alpha \sin\beta \cos\gamma + \sin\alpha \sin\gamma \\ r_{21} &= -\cos\beta \sin\gamma \\ r_{22} &= -\sin\alpha \sin\beta \sin\gamma + \cos\alpha \cos\gamma \\ r_{23} &= \cos\alpha \sin\beta \sin\gamma + \sin\alpha \cos\gamma \end{aligned} \qquad [12]$$

As in the two-dimensional problem, it is sought to minimize the mean square equation error

$$\begin{aligned} MSE = \frac{1}{N} \left|\left| \underline{u} - Mr_{11} (\underline{x} - \Delta x) - Mr_{12} (\underline{y} - \Delta y) - Mr_{13} \underline{z} \right|\right|^2 \\ + \frac{1}{N} \left|\left| \underline{v} - Mr_{21} (\underline{x} - \Delta x) - Mr_{22} (\underline{y} - \Delta y) - Mr_{23} \underline{z} \right|\right|^2 \end{aligned} \qquad [13]$$

where the ensemble of coordinate data are again arranged as vectors. Note that it is necessary to determine the unknown range vector $\underline{z}$, as well as the parameters Δx, Δy, M, α, β and γ.

Equating partial derivatives to zero leads to a set of transcendental algebraic equations to which a complete analytical solution was not, in fact, obtained. Instead, an algorithm of successive approximation was derived. It consists of a sequential application of analytical solutions to a set of subestimation (subminimization) problems.

The first iteration of the successive approximation algorithm consists of the following sequence of steps:

1) Estimate Δx, Δy, M, γ
   assuming $\alpha = \beta = \underline{z} = \emptyset$.

2) Estimate Δx, Δy, M, β, $\underline{z}$
   given γ from step 1
   assuming $\alpha = \emptyset$.

3) Estimate Δx, Δy, M, α, $\underline{z}$
   given γ from step 1
   assuming $\beta = \emptyset$.

4) If MSE of step 2 < MSE of step 3
   then estimate α, Δy
   given Δx, M, γ, β, $\underline{z}$ of step 2.
   If MSE of step 3 < MSE of step 2
   then estimate β, Δx
   given Δy, M, γ, α, $\underline{z}$ of step 3.

5) Estimate Δx, Δy, $\underline{z}$
given γ, α, β, M of step 4.

6) Estimate Δx, Δy, M, γ
given α, β, $\underline{z}$ of step 5.

Succeeding iterations consist of repetitions of steps 4, 5 and 6 with step 4 modified as follows:

4) Estimate α, Δy
given Δx, M, γ, β, $\underline{z}$ of step 6,
and
estimate β, Δx
given Δy, M, γ, α, $\underline{z}$ of step 6.

Step 1 is equivalent to the two-dimensional problem and requires no elaboration. The solution is given in Section 3.1 by equations 8.1 to 8.3. As will be seen later, this step generates a good estimate for roll angle γ in spite of the fact that the true transformation is three-dimensional.

Step 2 accepts the γ estimate and re-estimates the remaining parameters, including β and $\underline{z}$, with α still assumed to be zero. With the $\begin{bmatrix} u \\ v \end{bmatrix}$ data replaced by $\begin{bmatrix} \cos\gamma & \sin\gamma \\ -\sin\gamma & \cos\gamma \end{bmatrix}^{-1} \begin{bmatrix} u \\ v \end{bmatrix}$ the transformation can be expressed as

$$\begin{bmatrix} u \\ v \end{bmatrix} = M \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x-\Delta x \\ y-\Delta y \\ z \end{bmatrix} \qquad [14]$$

and the parameter estimates are

$$M = \frac{C_{vy}}{C_{yy}} \qquad [15.1]$$

$$\beta = \cos^{-1}\left(\frac{C_{ux} + \left[C_{ux}^{\,2} - \left[C_{xx} + C_{zz}\right] \bullet \left[C_{uu} - M^2 C_{zz} \right] \right]^{\frac{1}{2}}}{M \left[C_{xx} + C_{zz} \right]} \right) \qquad [15.2]$$

$$\underline{z} = \bar{z} - \frac{\underline{u} - \bar{u}}{M\sin\beta} + \frac{\underline{x} - \bar{x}}{\tan\beta} \qquad [15.3]$$

$$\Delta x = \bar{x} - \frac{\bar{u}}{M\cos\beta} - \bar{z}\tan\beta \qquad [15.4]$$

$$\Delta y = \bar{y} - \frac{\bar{v}}{M} \qquad [15.5]$$

The solution is not completely determined without specifying range statistics $\bar{z}$ and C_{zz} as well as the sign of the angle β. In other words, there are an infinite number of solutions that will optimally relate the reference grid to the search grid. For example, if C_{zz} is large, β will be small.

This nonuniqueness is a reflection of the loss of three-dimensional information in the process of imaging. Without additional external information, an image provides no depth information. Once the angle between two views of an object is known, two images can provide three-dimensional information. This, in fact, is the basis of stereometric depth perception (Refs. 8-9). In the tracking problem, however, the angle between two views is not known and hence unique determination of range is not possible. Even by using additional views and simultaneously matching multiple search grids (consecutive in time) to the reference grid one would not uniquely determine range without using angular information.

In application, values of $\bar{z}$ and C_{zz} are therefore assigned prior to the tracking mission. The sign of β can be determined on-line, however, by requiring that the computed range coordinates of the reference grid represent a convex surface. This is based on the assumption that the target is a reasonably shaped object whose front surface as seen in the reference image is essentially convex rather than concave. If the sign of β is wrong, the resulting range estimates describe a reflection of the true surface; that is, a convex surface becomes concave (and vice versa). The test for convexity consists in computing a coefficient of fit of the estimated range data to a

paraboloid. The sign of δ is taken to be the one that makes this convexity coefficient positive.

Step 3 is a parallel of step 2 but this time α is estimated and β is assumed to be zero. In the meantime, the results of step 2 are stored for later use. The transformation then is

$$\begin{bmatrix} u \\ v \end{bmatrix} = M \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \end{bmatrix} \begin{bmatrix} x-\Delta x \\ y-\Delta y \\ z \end{bmatrix} \qquad [16]$$

where, as in step 2, the (u,v) coordinates have been prerotated by a roll angle of γ. The estimates are

$$M = \frac{C_{ux}}{C_{xx}} \qquad [17.1]$$

$$\alpha = \cos^{-1}\left(\frac{C_{vy} + \left[C_{vy}^{\,2} - \left[C_{yy} + C_{zz} \right] \bullet \left[C_{vv} - M^2 C_{zz} \right] \right]^{\frac{1}{2}}}{M\,(C_{yy} + C_{zz})} \right) \qquad [17.2]$$

$$\underline{z} = \bar{z} + \frac{\underline{v} - \bar{v}}{M\sin\alpha} - \frac{\underline{y} - \bar{y}}{\tan\alpha} \qquad [17.3]$$

$$\Delta x = \bar{x} - \frac{\bar{u}}{M} \qquad [17.4]$$

$$\Delta y = \bar{y} - \frac{\bar{v}}{M\cos\alpha} + \bar{z}\tan\alpha \qquad [17.5]$$

The same values of $\bar{z}$ and C_{zz} as in step 2 are used, of course, and the sign of α is again determined by a convexity test on $\underline{z}$.

Step 4 compares the parallel estimates of steps 2 and 3 and accepts whichever yields the smaller mean square error. If the result of step 2 is chosen then step 4 will estimate α and Δy using estimates of Δx, M, β, γ, and $\underline{z}$ at the end of step 2. The procedure is to first compute $\underline{w}$ and prerotate (u,v,w) by angles γ and β. The problem is then simplified to that of determining the transformation

$$v = M\,[(y-\Delta y)\cos\alpha + z\sin\alpha] \qquad [18]$$

The parameter Δy is given by

$$\Delta y = \bar{y} - \frac{\bar{v}}{M\cos\alpha} + \bar{z}\tan\alpha \qquad [19.1]$$

and α is given implicitly by

$$C_{vz}\cos\alpha - C_{vy}\sin\alpha + M(C_{yy} - C_{zz})\sin\alpha\cos\alpha + M\,C_{yz}(\sin^2\alpha - \cos^2\alpha) = 0 \qquad [19.2]$$

which is solved using Newton's method.

If, instead, the estimate of step 3 is chosen, then step 4 will estimate β and Δx using estimates of Δy, M, α, γ, $\underline{z}$ at the end of step 3. The procedure is to prerotate (u,v) by γ and prerotate (x,y,z) by α. Then,it is required to determine the transformation

$$u = M(x-\Delta x)\cos\beta + z\sin\beta \qquad [20]$$

and the solution is given by

$$\Delta x = \bar{x} - \frac{\bar{u}}{M\cos\beta} - \bar{z}\tan\beta \qquad [21.1]$$

and

$$C_{uz}\cos\beta + C_{ux}\sin\beta - M(C_{xx} - C_{zz})\sin\beta\cos\beta + MC_{xz}(\sin^2\beta - \cos^2\beta) = 0 \qquad [21.2]$$

Step 5 accepts the magnification and angle estimates and updates the translation and range estimates. After minifying the (u,v) coordinates by a factor of M, the transformation is

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \end{bmatrix} \begin{bmatrix} x-\Delta x \\ y-\Delta y \\ z \end{bmatrix} \qquad [22]$$

where the matrix of r_{ij} is defined by eq. 12. The solution is

$$\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}^{-1} \begin{bmatrix} \bar{u} \\ \bar{v} \end{bmatrix} \qquad [23.1]$$

$$+ \frac{1}{r_{11}r_{22} - r_{12}r_{21}} \begin{bmatrix} r_{11}r_{13} - r_{12}r_{23} \\ r_{11}r_{23} - r_{21}r_{23} \end{bmatrix} \bar{z}$$

$$\underline{z} = \bar{z} + \frac{1}{r_{13}^2 + r_{23}^2} \Big(r_{13}(\underline{u} - \bar{u}) + r_{23}(\underline{v} - \bar{v}) - (r_{11}\, r_{13} + r_{21}\, r_{23})\,(\underline{x} - \bar{x}) - (r_{12}\, r_{13} + r_{22}\, r_{23})\,(\underline{y} - \bar{y}) \Big) \qquad [23.2]$$

Although the solution requires specification of only $\bar{z}$, consistency with the rest of the algorithm requires that the specification of C_{zz} be applied also. When this condition is added to the subestimation problem, the result is the same except that $\underline{z}$ is to be scaled to give the desired C_{zz}.

Step 6 performs an update of Δx, Δy, M, γ given α, β, $\underline{z}$. The solution is

$$\gamma = \tan^{-1}\left[\frac{C_{uz}\sin\alpha + C_{uy}\cos\alpha - C_{vx}\cos\beta - C_{vy}\sin\alpha\sin\beta + C_{vz}\cos\alpha\sin\beta}{C_{vz}\sin\alpha + C_{vy}\cos\alpha + C_{ux}\cos\beta + C_{uy}\sin\alpha\sin\beta - C_{uz}\cos\alpha\sin\beta}\right] \qquad [24.1]$$

$$M = \frac{\begin{aligned} & r_{11}\, C_{ux} + r_{12}\, C_{uy} + r_{13}\, C_{uz} \\ & + r_{21}\, C_{vx} + r_{22}\, C_{vy} + r_{23}\, C_{vz} \end{aligned}}{\begin{aligned} & s_{11}\, C_{xx} + s_{22}\, C_{yy} + s_{33}\, C_{zz} \\ & + (s_{12} + s_{21})\, C_{xy} + (s_{13} + S_{31})\, C_{xz} \\ & + (s_{23} + s_{32})\, C_{yz} \end{aligned}} \qquad [24.2]$$

$$\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} - \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}^{-1} \begin{bmatrix} \bar{u} \\ \bar{v} \end{bmatrix} + \frac{1}{r_{11}\, r_{22} - r_{12}\, r_{21}} \begin{bmatrix} r_{22}\, r_{13} - r_{12}\, r_{23} \\ r_{11}\, r_{23} - r_{21}\, r_{13} \end{bmatrix} \bar{z} \qquad [24.3]$$

where r_{ij} are as defined in eq. 12 and s_{ij} are elements of the matrix product

$$\begin{bmatrix} r_{11} & r_{21} \\ r_{12} & r_{22} \\ r_{13} & r_{23} \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \end{bmatrix}$$

In all the applications of this report, the algorithm shows quick convergence. The first iteration alone gives good results as illustrated in Fig. 8 and Tables I and II using a pyramid-type structure. Figure 8(a) shows a reference grid (broken lines) associated with the top view of a pyramid, on which is superposed a corresponding search grid (solid lines) where the pyramid has been translated, magnified, and rotated relative to the reference position. The transformation parameters are Δx = 1 picture element, Δy = 2 picture elements, M = 1.1, $\alpha = 10°$, $\beta = 20°$, $\gamma = 30°$. The magnification and rotations of this example are considerably larger than will be expected to occur in closed loop tracking.

The remainder of Fig. 8 illustrates the evolution of the parameter estimates by transforming the reference grid at each step using current estimates and superposing this on the search grid. The goodness of fit of the search grid and the transformed reference grid is a visual measure of the goodness of the estimates. Figures 8(b) to 8(f) show how the grids fit after each of the first 5 steps of the first iteration of the estimation algorithm. Note in Fig. 8(b) how step 1 has managed to generate a good estimate of the roll angle γ.

The results shown in Fig. 8 and Tables I and II were obtained with range statistics computed from exact values of $\underline{z}$. These statistics were $\bar{z} = 3.2$ and $C_{zz} = 20.48$. Since these values are generally not known to the tracker, other values were also tried, including an arbitrary choice of $\bar{z} = 0$ and $C_{zz} = \frac{1}{2}(C_{xx} + C_{yy})$, where C_{zz} is taken to be the average of C_{xx} and C_{yy}. The resulting parameter estimates were different, with the exception of magnification and roll angle whose

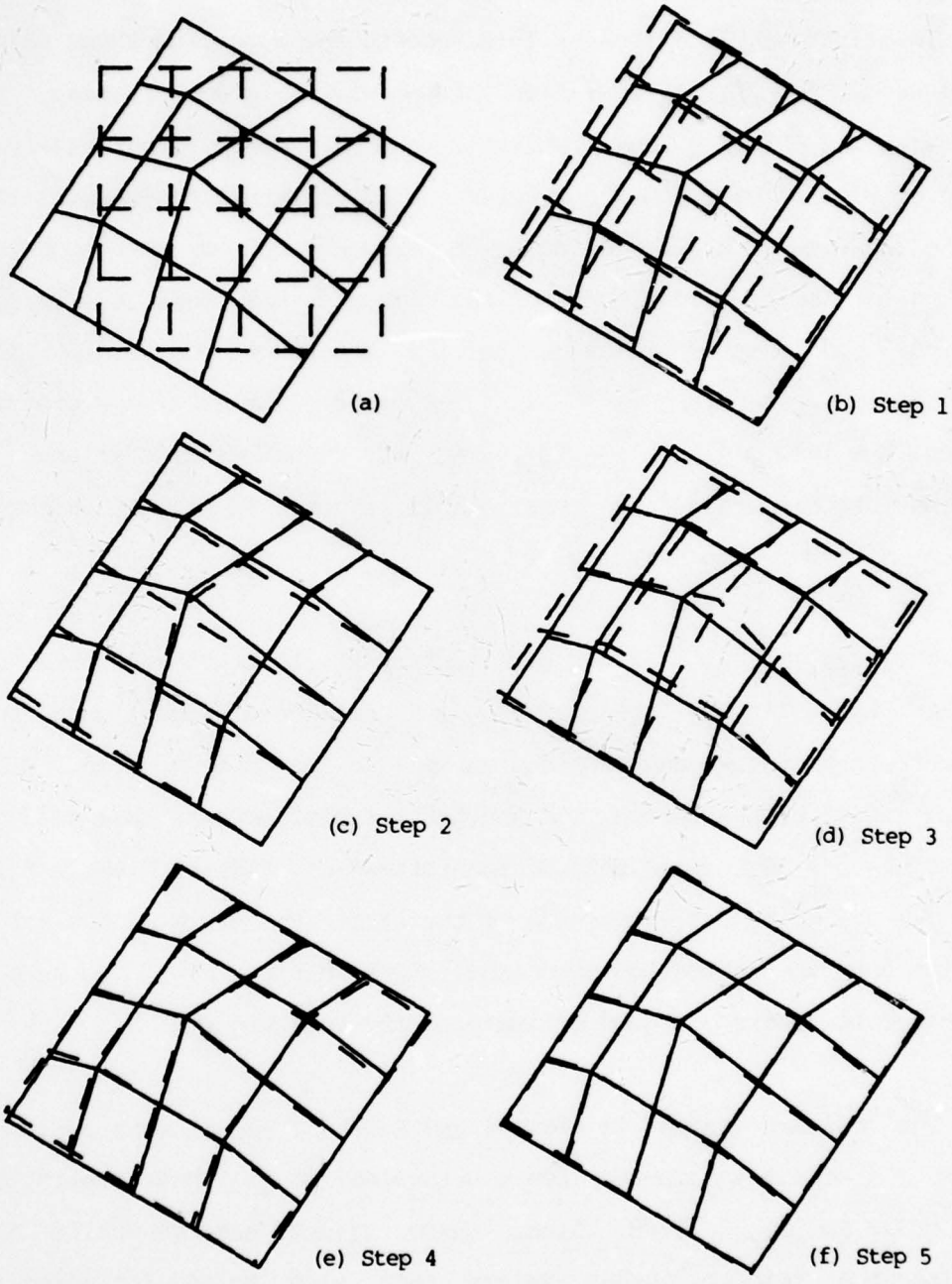


FIGURE 8 - Pyramid reference and search grids (a) and evolution of parameter estimates with first 5 steps of estimation algorithm (b) to (f)

dependence on $\bar{z}$ and C_{zz} appears to be only very weak. These solutions were equivalent, however, both visually and in terms of mean square error.

TABLE I

Evolution of parameter estimates and mean square error (MSE) for pyramid example

STEP #	Δx	Δy	M	α	β	γ	MSE
Initial	0.00	0.00	1.000	0.0	0.0	0.0	47.02
Step 1	2.17	1.54	1.059	0.0	0.0	31.8	1.99
Step 2	1.05	1.50	1.085	0.0	20.9	31.8	0.39
Step 3	2.22	2.73	1.033	18.5	0.0	31.8	1.51
Step 4	1.05	1.93	1.085	7.5	20.9	31.8	0.31
Step 5	0.96	1.93	1.085	7.5	20.9	31.8	0.10
Step 6	0.96	1.89	1.093	7.5	20.9	30.9	0.06
Exact	1.00	2.00	1.100	10.0	20.0	30.0	

TABLE II

Initial, estimated, and exact range parameters $\underline{z}$ for pyramid example

Initial $\underline{z}$

0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0

Estimated $\underline{z}$ (after step 6)

1.12	0.97	0.83	0.68	0.53
0.73	8.51	8.36	8.21	0.13
0.33	8.11	15.89	7.81	-0.27
-0.07	7.71	7.56	7.41	-0.67
-0.47	-0.62	-0.77	-0.92	-1.07

Exact $\underline{z}$

0	0	0	0	0
0	8	8	8	0
0	8	16	8	0
0	8	8	8	0
0	0	0	0	0

4.2 Small-Angle Simulation

In the two-dimensional case, it was shown that the original reference image can be used throughout the tracking mission. the live view of the target can always be made to coincide with the reference view by simple magnification and roll angle adjustment. Moreover, these correcting operations can be performed conveniently with a zoom lens and a gimbal.

This is no longer true when the target rotates about all 3 axes. Although the live scene can theoretically be rotated in 3 axes given a three-dimensional description of the target, a sufficiently distortion-free result would require a great deal of depth information, and hence, a prohibitive amount of computation for a tracking system. Moreover, the extent to which such an angular compensation can be carried out is limited, since it cannot regenerate parts of the target that rotate out of view.

Recourse to reference image updating is therefore taken. Constant scale and roll attitude are still controlled with a zoom lens and gimbal combination to maintain as much as possible the original optimally chosen configuration of target size and attitude. Rotations about perpendiculars to the line of sight (angles α and β) are estimated but not compensated for. Instead, a digitally scale- and roll-corrected section of the search scene is taken to replace the old reference image. The new reference, therefore, has the same scale and roll-orientation as the old reference; it differs only in the angles α and β.

The 3-D tracking algorithm can then be described as repetition of the following sequence:

1) Acquire the search scene.

2) Determine the search grid corresponding to the reference grid by image-matching.

3) Perform a three-dimensional grid analysis as described in Section 4.1 to estimate Δx, Δy, γ and M.

4) Send $\Delta x'$, $\Delta y'$ to sensor pitch and yaw control.

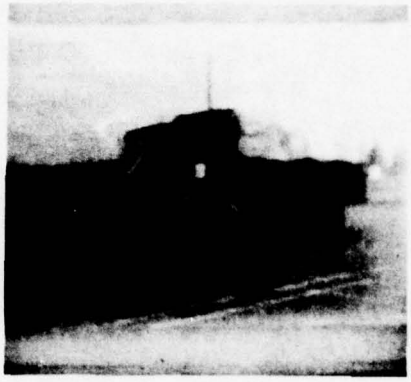
5) Send M to zoom lens control.

6) Send γ to sensor roll control.

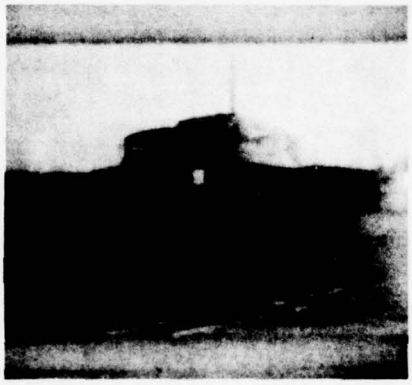
7) Replace the reference image by a section of search scene digitally centered at (Δx, Δy), roll-corrected by γ, and scaled by M,

where, as in the 2-D case, ($\Delta x'$, $\Delta y'$) are a prerotation of (Δx , Δy) by the total roll angle Γ.

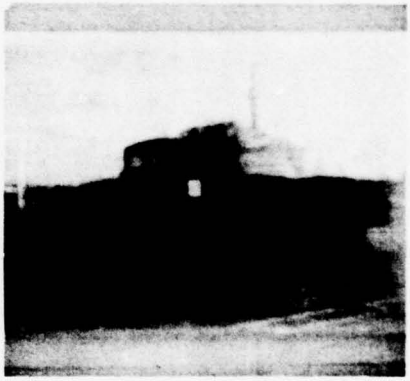
This algorithm was simulated on a series of polaroid photographs of a military truck shown in Fig. 9. After each photograph was taken, the camera was moved closer and to one side of the truck to produce an approximate 5% magnification and 3° azimuthal angle (β) between



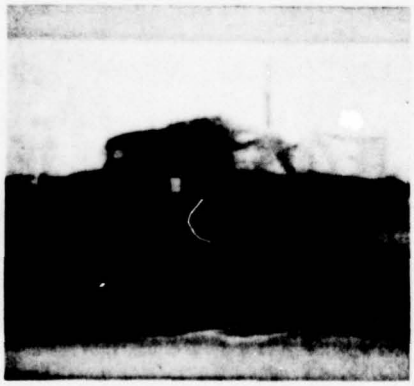
View 1



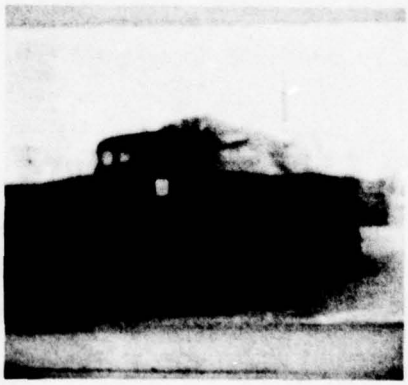
View 2



View 3

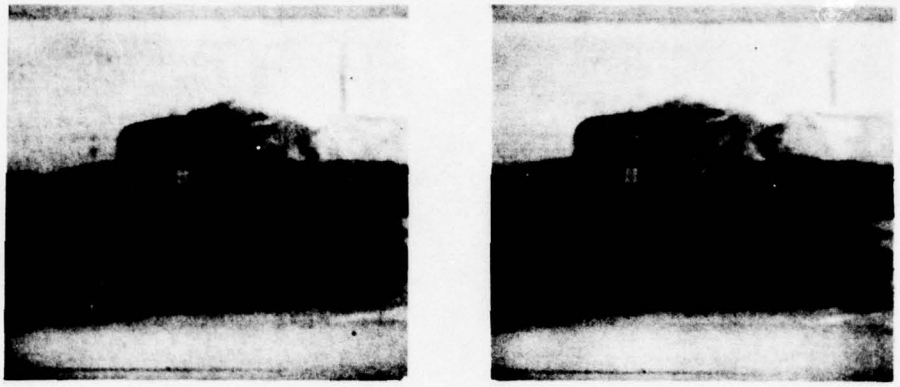


View 4



View 5

FIGURE 9 - Photographs of military truck used in tracking simulation



View 6 View 7

FIGURE 9 - (continued)

consecutive pictures. The photographs were digitized to a resolution of 128 x 128. View 1 was taken as the first reference, and Views 2 to 7 served as a set of six consecutive search scenes. The reference grid was defined by a 5 x 5 array of points with a spacing of 8 picture elements. Range parameters $\bar{z}$ and C_{zz} were preassigned values of $\bar{z} = 0$ and $C_{zz} = \frac{1}{2}(C_{xx} + C_{yy})$. The correcting action of the zoom lens was simulated by preminifying the search scene prior to correlation. Similarly, the roll correction was simulated by prerotation. Pitch and yaw corrections were not simulated since the truck was already sufficiently centered in all of the photographs.

To increase the speed, a two-step approach was used in determining the search grid. First, a wide search for a 32 x 32 element reference area using the fast sequential similarity detection algorithm provided the general location of the search grid. Then, a set of considerably narrower searches individually located the grid-points. In this secondary search, two refinements were made for a more accurate

pinpointing of the search grid points. One of these is the use of the cross-correlation algorithm rather than the SSDA. This is because the similarity function produced by the highly nonlinear SSDA is not meaningful for interpolation purposes. Moreover, the local centroid of the similarity function about the peak is not actually a good interpolation scheme. (Among other problems, the centroid moves with changing dc level.) Hence, as a second refinement, the centroid calculation is replaced by maximizing a local second-order polynomial approximation to the cross-correlation function. For simplicity, 2 one-dimensional interpolations along the x and y axes are made. The coordinates of the interpolated peak are then

$$x_{max} = x_p + \frac{1}{2}\left(\frac{F_w - F_e}{F_w - 2F_p + F_e}\right) \qquad [25.1]$$

$$y_{max} = y_p + \frac{1}{2}\left(\frac{F_s - F_n}{F_s - 2F_p + F_n}\right) \qquad [25.2]$$

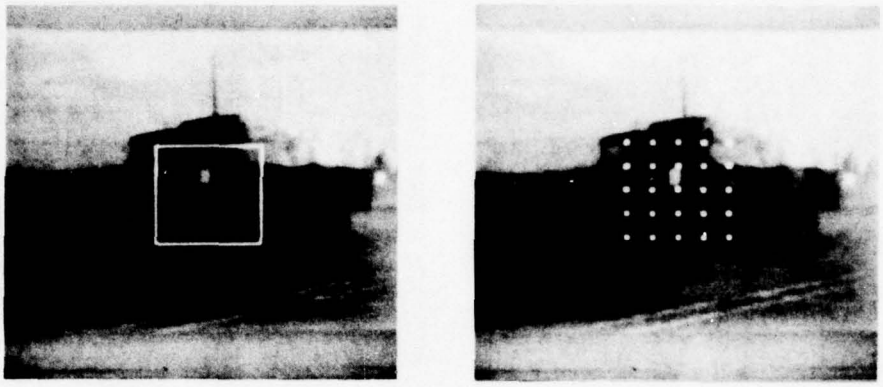
where F_w, F_e, F_p, F_s, F_n are neighbouring function values as shown below and F_p is the noninterpolated peak with integer coordinates (x_p, y_p).

$$\begin{array}{ccc} & F_n & \\ F_w & F_p & F_e \\ & F_s & \end{array}$$

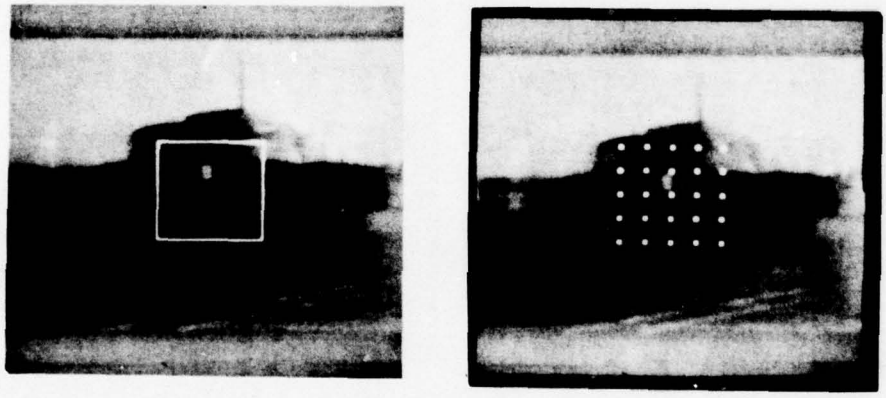
For comparison purposes, a parallel simulation was made to illustrate reference updating without benefit of grid analysis. In this run, which will be referred to as 1-D tracking, a single 32 x 32 reference was used, and only position information (Δx, Δy) determined the new reference. As in 3-D tracking, position was obtained by a wide SSDA search followed by a narrow cross-correlation search with second-order interpolation.

Parallel results of the 2 simulations are shown in Fig. 10. Reference image 1 is the first reference and is identical to View 1. Reference images 2 to 7 are updated references extracted from Views 2 to 7. The references of the 1-D simulation are shown on the left side of Fig. 10 with the 32 x 32 target outlined by a white square. The reference images of the 3-D simulation are shown on the right side of Fig. 10 superposed with the 5 x 5 array of reference grid points.

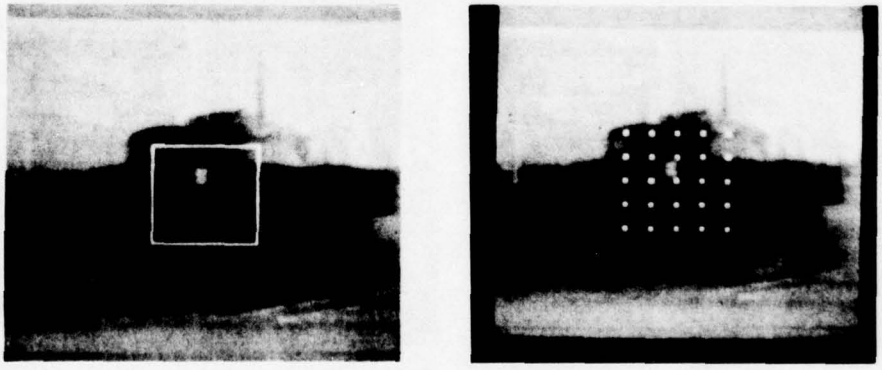
Inspection of Fig. 10 indicates that the 3-D tracking algorithm was very successful in maintaining original target scale and roll-orientation. Measurements on Views 1 and 7 indicate an overall magnification of 1.31. The product of the set of magnifications estimated by the 3-D algorithm is 1.311. Thus the 3-D tracker was able to effectively eliminate a 30% difference in scale factor spanning the series of photographs. For an onboard missile guidance application, this feature of maintaining original scale generally extends the useful range of the tracker. Without compensation, image expansion during target closure can lead to loss of target-defining detail, particularly when the target image overfills the field of view. This can be seen in the references of the 1-D simulation, where less and less of the truck is included in the target area. Apart from the door marking, the



Reference image 1

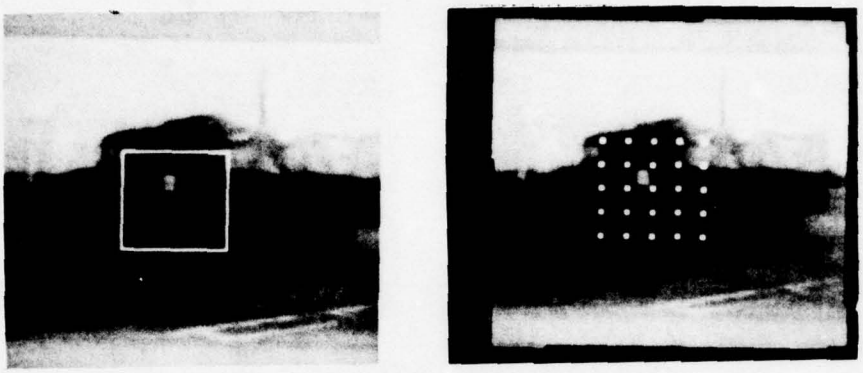


Reference image 2

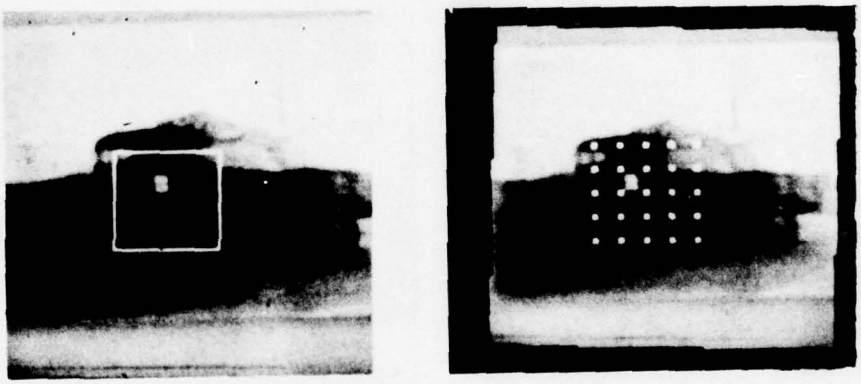


Reference image 3

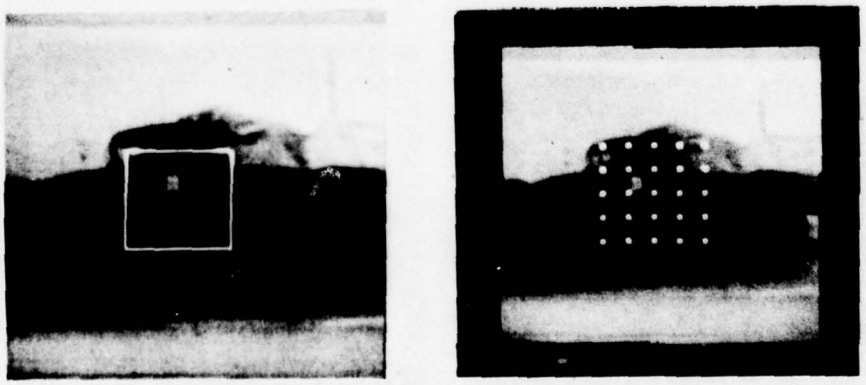
FIGURE 10 - Reference updates generated during tracking simulation with the truck target



Reference image 4

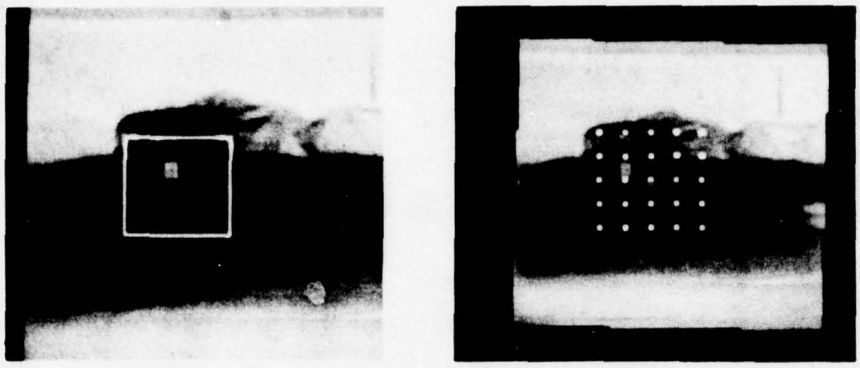


Reference image 5



Reference image 6

FIGURE 10 - (continued)



Reference image 7

FIGURE 10 - (continued)

outlined target of Reference image 7 retains very little picture detail to define a target. In Reference image 7 of the 3-D simulation, on the other hand, compensation for magnification resulted in retention of much more detail defining the cab of the truck (door, windshield, canvas, mounting step).

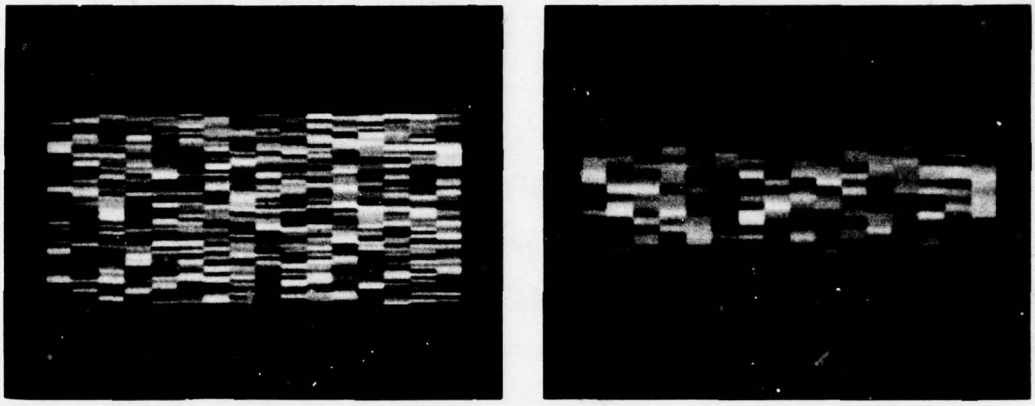
The other important distinction between the 2 simulation runs is in the positioning of new references. In the 1-D simulation, the system essentially tracked the high-contrast emblem on the door of the truck. It is evident in Fig. 10 that the target updates followed the door emblem (moving leftward) as the truck rotated. Eventually, if the emblem were to rotate out of view, the system would momentarily track the edge of the truck and finally move into the background with complete loss of the target.

In 3-D tracking, however, the system determines a three-dimensional description of the target. Therefore, it has the capability to discriminate between target translation and target rotation about perpendiculars to the line of sight. (The relatively unsophisticated 1-D and 2-D trackers both interpret rotation about perpendiculars to the line of sight as translation.) In Fig. 10, the resistance of the 3-D tracker to move left with the leftward rotation of the truck is quite evident when the grid position relative to the door emblem is noted.

4.3 Large-Angle Simulation

To test the performance of the 3-D tracker under large angles of rotation, a more convenient simulation was set up using a computer-generated cylinder as the target. The cylinder imagery was generated by mapping segments of a multishaded input plane (Fig. 11(a)) onto a semicylindrical surface, producing output cylinder images as shown in Fig. 11(b). The purpose of the random shading is to provide picture detail for tracking. The mapping incorporated a reflection coefficient consisting of the cosine of the angle between the viewing axis and the normal to the surface. Rotation about the x-axis (angle α) was performed simply by the appropriate selection of what segment of the input plane to map.

In this cylinder experiment, exact knowledge of the geometry together with nearly error-free search grids allowed a clear observation of the tracker's reaction to target rotation. It became apparent that exact tracking actually requires the use of the exact values of the



(a) Input plane

(b) Output cylinder image

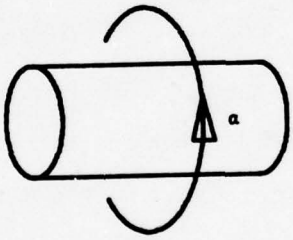


FIGURE 11 - Computer-generated cylinder image

range parameters $\bar{z}$ and C_{zz}. Although the value of C_{zz} has a weak effect on tracking accuracy, the value of $\bar{z}$ has a strong effect at large angles. This results from the fact that the mean parameter $\bar{z}$ defines what the tracker takes to be the geometric center of the target. If this point is not a valid center of rotation within the target, then the three-dimensional target motion as perceived by the tracker will include extraneous translations, with consequent positional update errors. If $\bar{z}$ is made too small, the system will tend to undercorrect; that is, a reduced ability to maintain its aim on target center and it will have a tendency to follow the original reference image as it rotates out of view. If $\bar{z}$ is made too large, the system will overcorrect; it will move its aim too far in the direction opposite to the target rotation.

All this does not mean that good tracking cannot be achieved without exact knowledge of $\bar{z}$ and C_{zz}. If the target is a reasonably regularly shaped object, then approximate values of $\bar{z}$ and C_{zz} that place the center of rotation almost anywhere within the target will allow 3-D tracking to a certain extent. The more the target rotates away from the original reference attitude, the more accurate an estimate of $\bar{z}$ is required to maintain lock on target.

Based on the preceding argument, approximate values of $\bar{z}$ and C_{zz} are obtained by modelling the target as a spherical object. The points of the 5 x 5 reference grid are therefore modelled as points on the surface of a sphere. Since the reference grid is supposed to more or less represent the target, the modelled grid is made to inscribe the outline of the sphere as diagrammed in Fig. 12. In this experiment, the grid spacing is 8 picture, elements. Therefore the radius of the sphere

is 22.6 picture elements and the range statistics are $\bar{z}$ = 14.4 and C_{zz} = 48.0. The corresponding range statistics of the same grid on the cylinder (whose radius is 24 picture elements) are $\bar{z}$ = 21.0 and C_{zz} = 6.7. In this case, the spherical target model leads to a significantly underestimated $\bar{z}$ and very overestimated C_{zz}.

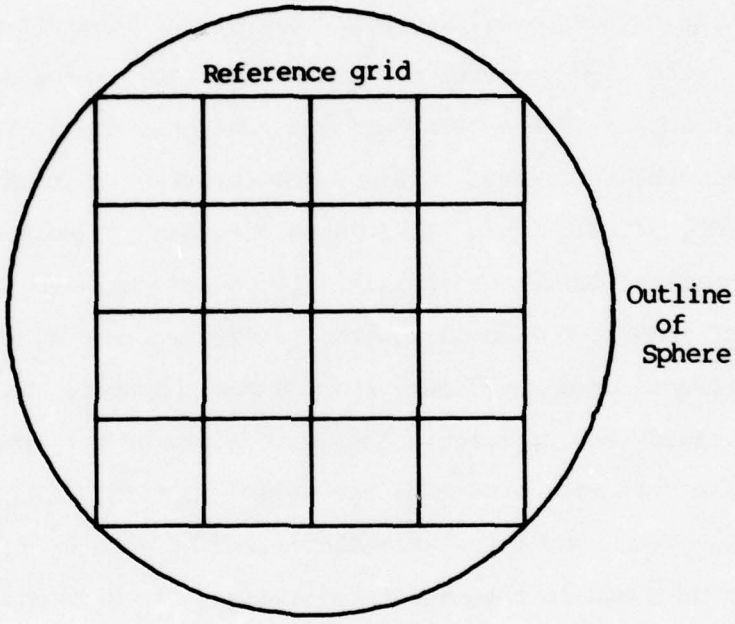
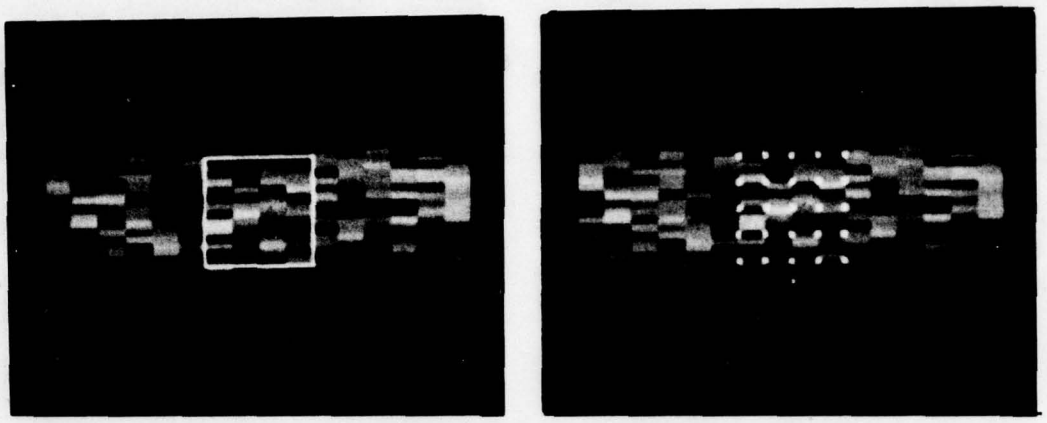


FIGURE 12 - Geometry of the spherical target model

For the cylinder tracking simulation, a series of views was generated with increasing angles of cylinder rotation about the x-axis. The relative angle of rotation between consecutive views was 5°. As in the simulation with the series of truck photographs, both the 1-D and 3-D tracking algorithms were applied. The only difference here was the use of range parameters given by the spherical target model.

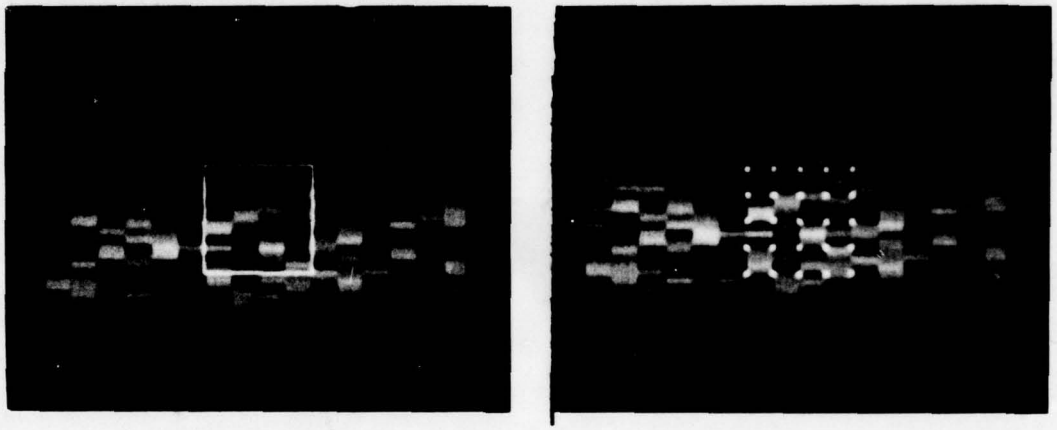
Results of the parallel 1-D and 3-D simulations are shown in Fig. 13 (again with 1-D references on the left and 3-D references on the right). Although a reference update was made at every 5° interval, Fig. 13 shows only the references obtained from views corresponding to α = 0°, 30°, 60°, 90° and 120°. It is apparent that the underestimation of $\bar{z}$ by the spherical target model initially caused the 3-D algorithm to undercorrect and tend to follow the rotation of the cylinder. As the updated reference grid moved away from center, however, the true value of $\bar{z}$ decreased and approached the value given by the spherical target model. While this was happening, the upward motion of the reference grid slowed down and the system stabilized at a point on the cylinder where the true $\bar{z}$ equals the estimated value of $\bar{z}$ used by the 3-D algorithm. This stabilization occurred in this case at α = 45°. From that point on, the system's aim remained stable on the cylinder while the cylinder was rotated further to α = 120°.

The 1-D tracker, on the other hand, tried to follow the cylinder's image rather than track it as a three-dimensional object. Consequently, its aimpoint rolled off target as the cylinder rotated. At α = 60° the aimpoint was approximately on the edge of the cylinder. The 32x32 (updated) reference image was then half cylinder and half



Reference image 1

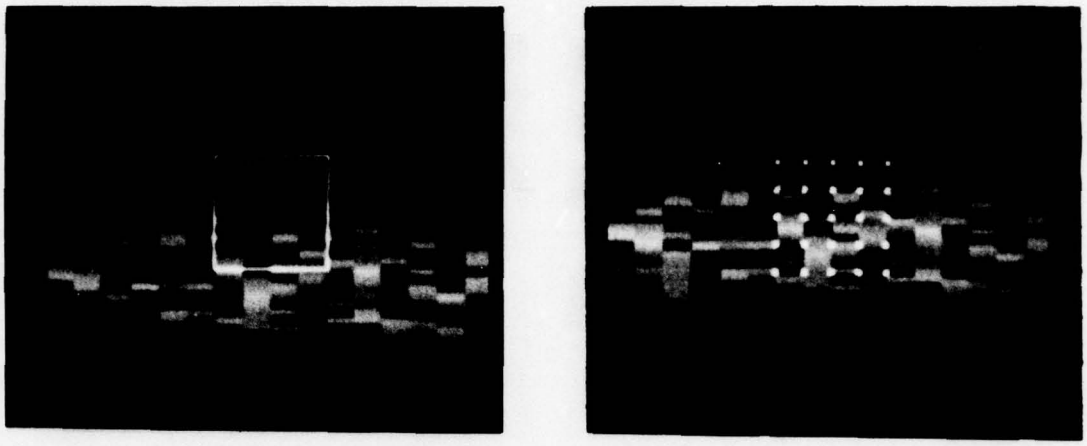
$\alpha = 0^{\circ}$



Reference image 7

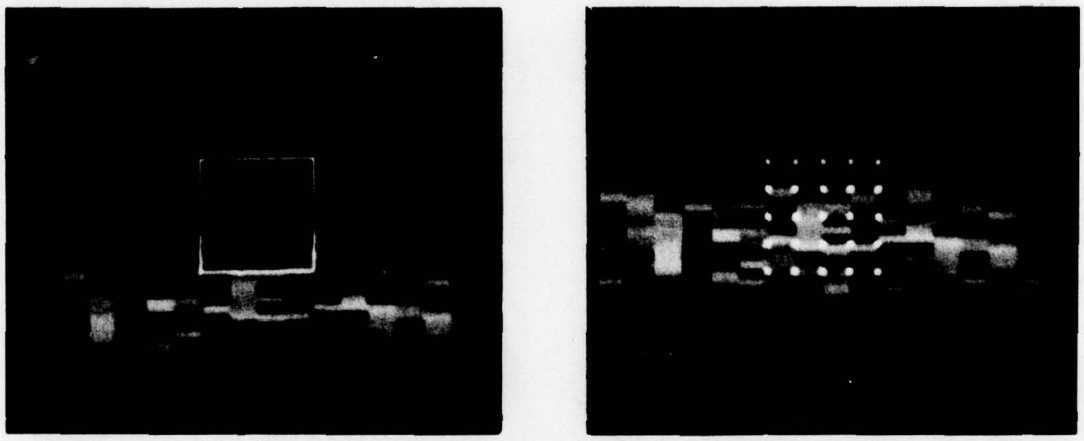
$\alpha = 30^{\circ}$

FIGURE 13 - Reference updates generated during tracking simulation with the cylinder target



Reference image 13

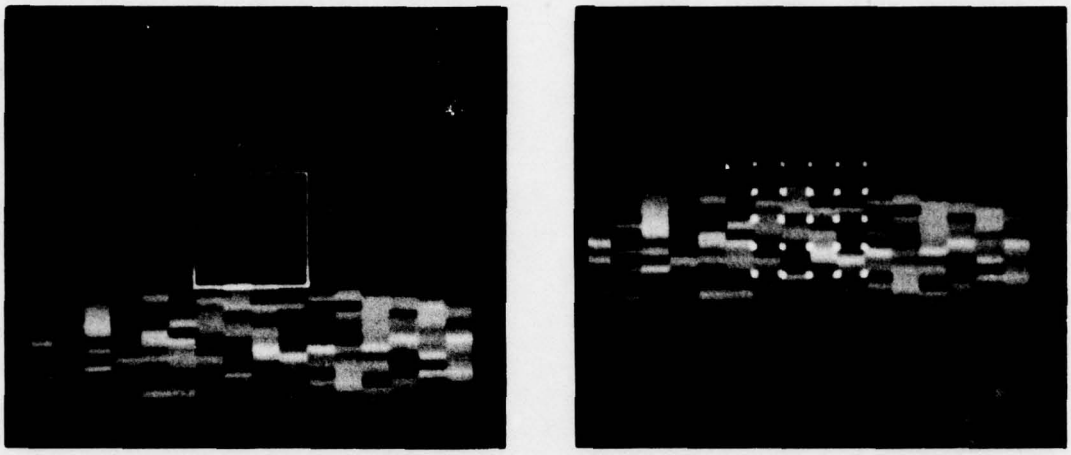
$\alpha = 60^{\circ}$



Reference image 19

$\alpha = 90^{\circ}$

FIGURE 13 - (continued)



Reference image 25

$\alpha = 120^{\circ}$

FIGURE 13 - (continued)

background. At this point, if the cylinder had been moving across a non-empty background, the background would quickly have taken over as the target. The tracker would then have been left fixed on a point in the background and it would have lost the target entirely. In this simulation, the black background did not exert any pull on the tracker's aim, and thus the aiming point moved relatively slowly away from the cylinder even when the reference image contained very little of the cylinder. Figure 14 compares the performances of the 1-D and 3-D algorithms in terms of aiming point deviation from the center of the cylinder. The stabilization of the 3-D tracker after about 45° of cylinder rotation is clearly evident.

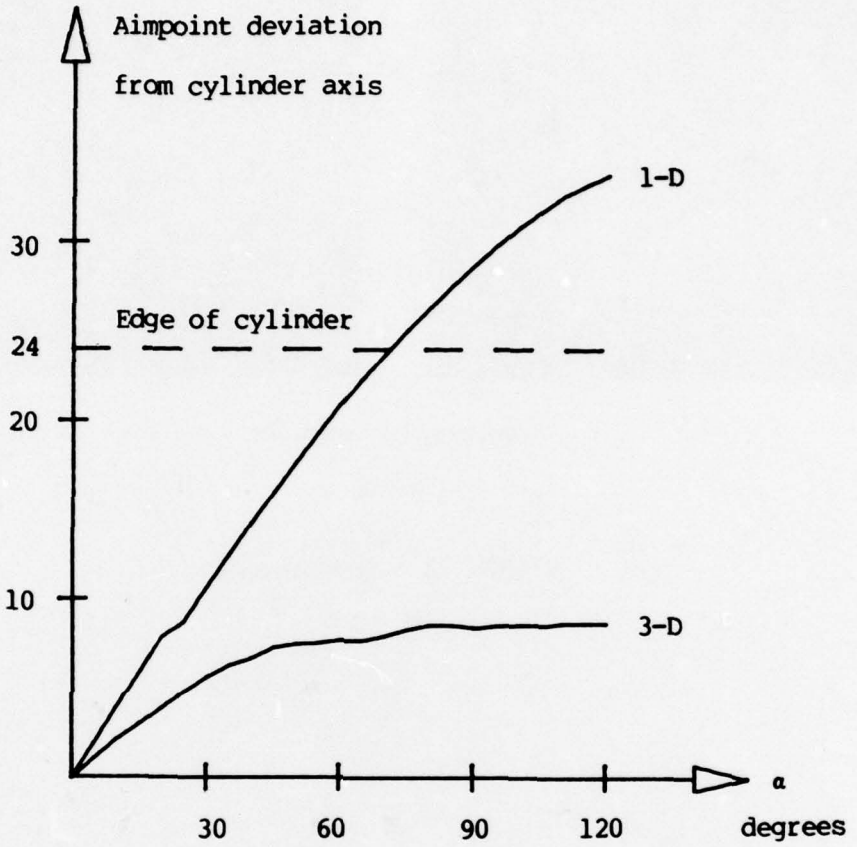


FIGURE 14 - Aimpoint deviation produced by 1-D and 3-D tracking algorithm, as a function of rotation angle α

Although the 2-D tracker is a non-updating algorithm intended to be restricted to the two-dimensional case, it is instructive to note the consequence of its application to the three-dimensional problem; that is, an updating version of the 2-D tracker applied to a target rotating in three axes. In comparison to the 1-D tracker, the 2-D tracker is more sophisticated in that it can detect and compensate for magnification and rotation about the line of sight (roll). It is not capable of perceiving the target as a three-dimensional object, and it is, therefore, still restricted to following images. Like the 1-D tracker, its aim will tend to roll off-target as the original reference view rotates out of sight. Although the 2-D tracker does not detect rotation about perpendiculars to the line of sight, it does, however, misinterpret the distorting effect such rotation has on search grids. As an example, without three-dimensional intelligence, the 2-D tracker erroneously interprets a vertically squashed search grid from the rotating cylinder as an overall reduction in the size of the target. It therefore instructs the zoom lens control to perform a compensatory magnification. Thus the reaction of the 2-D tracker to three-dimensional target rotation includes magnifying the target image. This behavior is evident in the cylinder simulation in the first step of the three-dimensional coordinate estimation algorithm (this first step is identical to the two-dimensional coordinate estimation). By the time the cylinder has rotated to 120° in the simulation, the initial two-dimensional algorithm interpreted the 5° step in rotation about the x-axis as a reduction in scale by a factor of 0.975. (This would have caused a magnification by a factor of 1.0256.)

5.0 APPLICATION TO MISSILE GUIDANCE

Implementations of the 2-D and 3-D tracking algorithms would produce very different missile guidance systems. The computational workload of the 2-D algorithm can be handled by a medium performance microprocessor supported by high-speed digital image-matching hardware. The 2-D algorithm is therefore suitable for a missile-borne tracking system. This is a fire-and-forget missile system, in which the missile is totally autonomous after launch. During target closure, the 2-D tracker performs its own magnification and roll angle corrections. The lack of compensation for rotation in the other 2 axes restricts usage of the 2-D system generally to fixed land targets, preferably in the air-to-ground configuration.

Like the 2-D tracking algorithm, the 3-D algorithm controls scale and roll angle (via zoom lens and roll gimbal) according to information obtained from grid analysis. It also uses this information, however, to update the reference image by periodically taking a geometrically corrected search scene as a new reference. The geometric precorrection of the search scene, based on grid analysis data, consists of a simultaneous translation (horizontal and vertical), change of scale, and rotation (roll angle). This transformation, involving interpolation in two dimensions, is quite a massive task of floating point computation. The required processing power would be bulky and rather expensive for an expendable system. Because of this, the 3-D tracking algorithm is probably not feasible as a missile-borne guidance system.

A nonexpendable application for the 3-D algorithm, in which the size and weight restrictions are relaxed, is an automatic airborne laser designator. In this configuration, the pilot defines the target, hands off control to the automatic designator and launches the missile. From that point on, the system is autonomous, leaving the pilot free to maneuver without the burden of directing the missile. The designator tracks it and maintains a laser beam on the target. The task of the missile itself is simply to fly toward the laser light reflected by the target. This greatly reduces the cost of the missile system since the sophisticated tracking system is shared by a large number of relatively inexpensive missiles. Another advantage is that the pilot can, at any time, take control to redefine the target to correct a break-lock or an initial aiming error. The disadvantage of such a system is that although it is a short-term fire-and-forget system freeing the pilot from the designation task, it is not long-term fire-and-forget in the sense that the pilot is not free to escape outside the designator's range until termination of the missile's flight.

Although 3-D tracking is required against maneuvering targets, it is also essential against fixed targets when configured as an airborne designator. This is because a target's aspect will vary as the aircraft maneuvers around the target area. Consider, as a simple example, an attack mission in which a 700-mi/h aircraft in straight and level flight launches a 1000-mi/h missile toward a fixed land target at a 45° lookdown angle. At missile impact, the aircraft would be directly above the target. This is a change in target aspect with respect to the aircraft. Accurate tracking and designation throughout missile flight would therefore require 3-D tracking.

Apart from noise and atmospheric conditions, range of operation is determined by the optics and the physical size of the target. If the format of the imaging sensor has a horizontal dimension H, and the required horizontal field of view angle at launch is θ, then the required telephoto focal length of the zoom lens is given by

$$f_{tele} = \frac{H}{2 \tan\left[\frac{\theta}{2}\right]}$$

Small focal lengths are highly desirable for compact packaging and convenient gimballing of the optics. For a given field of view, a small focal length requires a small image format. In this application, the small format of charge-coupled device (CCD) imaging sensors (Ref. 10) may be used to advantage. As an example, the image format of a commercially available 512 x 320 element CCD imager is 7.30 mm x 9.75 mm (H = 9.75 mm). A telephoto focal length of 75 mm would give a total field of view of about 7.5 degrees. If the target occupies 32 elements of the 320-element horizontal scan, then it would occupy a field of view of 0.75 degree. This is the angle subtended by a 50-m target at a range of about 4 kilometres.

As the missile closes in on target, the zoom lens must be directed toward the wide-angle position (decreasing focal length). The greater the zoom ratio, the closer the missile can approach the target before it can no longer maintain original target size in its field of view. With a zoom ratio of 10, in the example just described, the missile could close in to within 400 m of the target without losing the original scale.

6.0 CONCLUSION

Two algorithms for determining relative image magnification and rotation have been presented. The information provided by these algorithms can be used by an area correlation tracker to perform zoom and roll corrections so as to maintain original target scale and roll attitude. This allows a missile with an onboard area correlator to approach a fixed target without updating the reference image, thus avoiding updating errors.

The algorithms operate by defining a grid of points on the reference target image and determining the corresponding grid of points in the search scene by matching image segments. The geometric relationship between the 2 images (translation, magnification, rotation) is given by a least squares estimate of the coordinate transformation between the reference grid and the search grid.

The 2-D algorithm assumes that the target, as viewed by the sensor, does not rotate about perpendiculars to the line of sight. This is the case of a missile flying along the line of sight toward a fixed target. In this case, the coordinate transformation between grids can be described in 2 dimensions and the estimate of the transformation is given in closed form. A simulation with an aerial infrared photograph of a highway illustrates the 2-D algorithm's capability.

When there are rotations about the line of sight, as in the case of maneuvering targets, the geometric relationship between reference and search grids must be described in 3 dimensions. Since images are

two-dimensional and do not provide direct depth information, the 3-D algorithm estimates not only the three-dimensional coordinate transformation, but the missing depth information as well. Not all of the depth information is recoverable, however, and the solution is not unique. A priori estimates of the mean and variance of grid point depth must be provided to determine the solution. The mean defines for the tracker the target's center of rotation, and the variance gives the tracker a measure of the overall depth of the target. The solution is given as an algorithm of successive approximation.

As in the 2-D algorithm, the 3-D tracker corrects for magnification and roll via a zoom lens and roll gimbal. No compensatory action can be taken, however, to eliminate effects of rotation about perpendiculars to the line of sight. The 3-D tracker therefore updates the reference by selecting a scale-and roll-corrected search scene as a new reference image. Although the target aspect may change, the scale and roll attitude of the reference image remains the same.

A tracking system equipped with 3-D processing acquires an important distinction over other area correlation systems. An area correlator without grid processing is essentially a point-tracker; that is, it basically tracks a single point on the target. With 2-D processing, the system tracks a collection of target-points which are assumed to lie in a plane perpendicular to the line of sight. Such a system can maintain a constant target scale and roll attitude in its field of view and, thus, can be considered to be tracking an image of the target rather than simply a point. Like the point-tracker, however, the image tracker would fail as the original view of the target rotates

out of view. Only 3-D processing allows the system to track a target as a three-dimensional object and, thus, to maintain aim at or close to target center even after the original view has disappeared.

Two simulations were performed to test the 3-D tracking algorithm. In a simulation utilizing a series of photographs of a military truck, the 3-D algorithm proved successful at maintaining correct aim and constant target image scale in the face of magnification and three-dimensional rotation. The resistance of the 3-D tracker's aim point to roll offtarget with target rotation was further tested on a series of computer-generated and computer-rotated cylinder images. The aim remained close to center for cylinder rotation as great as 120 degrees. It was shown that exact tracking with zero wander requires exact knowledge of depth mean and variance. If the depth mean is underestimated, the tracker's aim will initially follow the rotation of the target, but will stabilize, however, at a point where actual depth mean equals the assumed depth mean.

The 2-D tracking algorithm, implemented in a microprocessor supported by image-matching hardware, is suitable as the basis of a missile-borne fire-and-forget guidance system. The most appropriate application would be air-to-ground against fixed targets. The 3-D algorithm is not likely to be suitable for a missile-borne guidance system because of the heavy computation involved in the geometricallycorrected reference updating scheme. A more promising role for the 3-D tracker would be an automatic airborne laser designator. Because of target aspect variations caused by aircraft flight, 3-D tracking is advantageous in this application.

7.0 REFERENCES

1. Rosenfeld, A., "Picture Processing by Computer", Academic Press, New York, 1969.

2. Andrews, H. C., "Computer Techniques in Image Processing", Academic Press, New York, 1970.

3. Anuta, P. E., "Spatial Registration of Multispectral and Multitemporal Digital Imagery Using Fast Fourier Transform Techniques", IEEE Transactions on Geoscience Electronics, Vol. GE-8, pp. 353-368, October 1970.

4. Bailey, H. H., Blackwell, F. W., Lowery, C. L., and Ratkovic, J. A., "*Image Correlation: Part I. Simulation and Analysis*", R-2057/1-PR Rand Corporation, Santa Monica, California, November 1976, UNCLASSIFIED

5. Wessely, H. W., "Image Correlation: Part II. Theoretical Basis", R-2057/2-PR Rand Corporation, Santa Monica, California, November 1976, UNCLASSIFIED

6. Barnea, D. I., and Silverman, H. F., "A Class of Algorithms for Fast Digital Image Registration", IEEE Transactions on Computers, Vol. C-21, pp. 179-186, February 1972.

7. Munteanu, C., "Digital Area Correlation Tracking by Sequential Similarity Detection", DREV R-4097/77, November 1977, UNCLASSIFIED

8. Levine, M. D., O'Handley, D. A., and Yagi, G. M., "Computer Determination of Depth Maps", Computer Graphics and Image Processing, Vol. 2, pp. 131-150, 1973.

9. Nevatia, R., "Depth Measurement by Motion Stereo", Computer Graphics and Image Processing, Vol. 5, pp. 203-214, 1976.

10. Sequin, C. H., and Tompsett, M. F., "Charge Transfer Devices", Academic Press, New York, 1975.

CRDV R-4140/79 (NON CLASSIFIE)

Bureau - Recherche et Développement, MDN, Canada.
CRDV, C.P. 880, Courcelette, Qué. G0A 1R0

"Poursuite par la corrélation avec la compensation automatique pour l'agrandissement et la rotation"
par C. Munteanu

La poursuite d'une cible par la corrélation d'images consiste à corréler une image étalon de la cible avec une scène prise sur le vif. Le grossissement et les rotations de la cible diminuent considérablement les performances de l'algorithme de corrélation. Le présent rapport décrit deux algorithmes qui permettent à la tête chercheuse de compenser l'effet du grossissement et des rotations à l'aide d'une lentille zoom et d'un système de balanciers. Un premier algorithme, applicable à des cibles fixes, évalue l'angle de rotation par rapport à la ligne de visée (roulis). Quant au second, il détermine l'angle de rotation selon les trois axes, ce qui est nécessaire pour une cible en mouvement. Une simulation réalisée à partir de photographies et d'images artificielles démontre l'efficacité de ces deux algorithmes. (NC)

CRDV R-4140/79 (NON CLASSIFIE)

Bureau - Recherche et Développement, MDN, Canada.
CRDV, C.P. 880, Courcelette, Qué. G0A 1R0

"Poursuite par la corrélation avec la compensation automatique pour l'agrandissement et la rotation"
par C. Munteanu

La poursuite d'une cible par la corrélation d'images consiste à corréler une image étalon de la cible avec une scène prise sur le vif. Le grossissement et les rotations de la cible diminuent considérablement les performances de l'algorithme de corrélation. Le présent rapport décrit deux algorithmes qui permettent à la tête chercheuse de compenser l'effet du grossissement et des rotations à l'aide d'une lentille zoom et d'un système de balanciers. Un premier algorithme, applicable à des cibles fixes, évalue l'angle de rotation par rapport à la ligne de visée (roulis). Quant au second, il détermine l'angle de rotation selon les trois axes, ce qui est nécessaire pour une cible en mouvement. Une simulation réalisée à partir de photographies et d'images artificielles démontre l'efficacité de ces deux algorithmes. (NC)

CRDV R-4140/79 (NON CLASSIFIE)

Bureau - Recherche et Développement, MDN, Canada.
CRDV, C.P. 880, Courcelette, Qué. G0A 1R0

"Poursuite par la corrélation avec la compensation automatique pour l'agrandissement et la rotation"
par C. Munteanu

La poursuite d'une cible par la corrélation d'images consiste à corréler une image étalon de la cible avec une scène prise sur le vif. Le grossissement et les rotations de la cible diminuent considérablement les performances de l'algorithme de corrélation. Le présent rapport décrit deux algorithmes qui permettent à la tête chercheuse de compenser l'effet du grossissement et des rotations à l'aide d'une lentille zoom et d'un système de balanciers. Un premier algorithme, applicable à des cibles fixes, évalue l'angle de rotation par rapport à la ligne de visée (roulis). Quant au second, il détermine l'angle de rotation selon les trois axes, ce qui est nécessaire pour une cible en mouvement. Une simulation réalisée à partir de photographies et d'images artificielles démontre l'efficacité de ces deux algorithmes. (NC)

CRDV R-4140/79 (NON CLASSIFIE)

Bureau - Recherche et Développement, MDN, Canada.
CRDV, C.P. 880, Courcelette, Qué. G0A 1R0

"Poursuite par la corrélation avec la compensation automatique pour l'agrandissement et la rotation"
par C. Munteanu

La poursuite d'une cible par la corrélation d'images consiste à corréler une image étalon de la cible avec une scène prise sur le vif. Le grossissement et les rotations de la cible diminuent considérablement les performances de l'algorithme de corrélation. Le présent rapport décrit deux algorithmes qui permettent à la tête chercheuse de compenser l'effet du grossissement et des rotations à l'aide d'une lentille zoom et d'un système de balanciers. Un premier algorithme, applicable à des cibles fixes, évalue l'angle de rotation par rapport à la ligne de visée (roulis). Quant au second, il détermine l'angle de rotation selon les trois axes, ce qui est nécessaire pour une cible en mouvement. Une simulation réalisée à partir de photographies et d'images artificielles démontre l'efficacité de ces deux algorithmes. (NC)

DREV R-4140/79 (UNCLASSIFIED)

Research and Development Branch, DND, Canada.
DREV, P.O. Box 880, Courcelette, Que. G0A 1R0

"Area correlation tracking with automatic compensation for magnification and rotation"
by C. Munteanu

Target tracking by area correlation involves matching a stored reference image of the target to live search scenery. The performance of the area correlator can be severely degraded by target magnification and three-axis rotation. This report presents 2 algorithms that estimate magnification and rotation angles, and allow the tracker to take compensating action via zoom lens and roll gimbal. The first algorithm, applicable to fixed targets, estimates rotation about the line of sight only (roll angle). The second algorithm estimates rotation in 3 axes, as required for maneuvering targets. The presentation of both algorithms is supported by simulation using photographs and artificial images. (U)

DREV R-4140/79 (UNCLASSIFIED)

Research and Development Branch, DND, Canada.
DREV, P.O. Box 880, Courcelette, Que. G0A 1R0

"Area correlation tracking with automatic compensation for magnification and rotation"
by C. Munteanu

Target tracking by area correlation involves matching a stored reference image of the target to live search scenery. The performance of the area correlator can be severely degraded by target magnification and three-axis rotation. This report presents 2 algorithms that estimate magnification and rotation angles, and allow the tracker to take compensating action via zoom lens and roll gimbal. The first algorithm, applicable to fixed targets, estimates rotation about the line of sight only (roll angle). The second algorithm estimates rotation in 3 axes, as required for maneuvering targets. The presentation of both algorithms is supported by simulation using photographs and artificial images. (U)

DREV R-4140/79 (UNCLASSIFIED)

Research and Development Branch, DND, Canada.
DREV, P.O. Box 880, Courcelette, Que. G0A 1R0

"Area correlation tracking with automatic compensation for magnification and rotation"
by C. Munteanu

Target tracking by area correlation involves matching a stored reference image of the target to live search scenery. The performance of the area correlator can be severely degraded by target magnification and three-axis rotation. This report presents 2 algorithms that estimate magnification and rotation angles, and allow the tracker to take compensating action via zoom lens and roll gimbal. The first algorithm, applicable to fixed targets, estimates rotation about the line of sight only (roll angle). The second algorithm estimates rotation in 3 axes, as required for maneuvering targets. The presentation of both algorithms is supported by simulation using photographs and artificial images. (U)

DREV R-4140/79 (UNCLASSIFIED)

Research and Development Branch, DND, Canada.
DREV, P.O. Box 880, Courcelette, Que. G0A 1R0

"Area correlation tracking with automatic compensation for magnification and rotation"
by C. Munteanu

Target tracking by area correlation involves matching a stored reference image of the target to live search scenery. The performance of the area correlator can be severely degraded by target magnification and three-axis rotation. This report presents 2 algorithms that estimate magnification and rotation angles, and allow the tracker to take compensating action via zoom lens and roll gimbal. The first algorithm, applicable to fixed targets, estimates rotation about the line of sight only (roll angle). The second algorithm estimates rotation in 3 axes, as required for maneuvering targets. The presentation of both algorithms is supported by simulation using photographs and artificial images. (U)